Prompt Photon Production in Photoproduction at HERA

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr.sc.nat.)

vergelegt der
Mathematisch-naturwissenschaflichen Fakultät
der
Universität Zürich
von

Krzysztof Nowak
aus
Polen

Promotionskomitee
Prof. Dr. Ulrich Straumann (Vorsitz)
Dr. Katharina Müller

Zürich, 2009

Promotionskomitee
Prof. Dr. Ulrich Straumann
Dr. Katharina Müller

Abstract

This thesis presents measurement of the production of prompt photons in photoproduction with the H1 experiment at HERA. The analysis is based on the data taken in the years 2004-2007, with a total integrated luminosity of 340 pb^{-1}. The main difficulty of the measurement comes from the high background of neutral mesons decaying into photons. It is accounted for with the help of multivariate analysis. Prompt photon cross sections are measured with the low negative four-momentum transfer squared $Q^2 < 1\,\mathrm{GeV}^2$ and in the inelasticity range $0.1 < y < 0.7$ for photons with a transverse energy $6 < E_T^\gamma < 15\,\mathrm{GeV}$ and in the pseudorapidity range $-1.0 < \eta^\gamma < 2.4$ as a function of photons transverse energy and its pseudorapidity.

Cross sections for prompt photon events with an additional hadronic jet are measured as a function of the transverse energy and pseudorapidity of the jet and of the momentum fractions x_γ and x_p of the incident photon and proton carried by the constituents participating in the hard scattering process. Additionally, the transverse correlation between the photon and the jet is studied. The results are compared with predictions of a next-to-leading order calculation and a calculation based on the k_T factorisation approach. Neither of calculations is able to describe all the aspects of the measurement.

Kurzfassung

In dieser Arbeit wird die Produktion von prompten Photonen in Photoproduktion mit dem H1 Detektor bei HERA vorgestellt. Die Analyse basiert auf den Daten der Datennahmeperiode 2004-2007 mit einer integrierten Luminosität von 340 pb^{-1}. Die Messung wird erschwert durch den hohen Anteil von Untergrundereignissen vom Zerfall von neutralen Mesonen. Dieser wird mit Hilfe einer multivariaten Analyse bestimmt. Der Wirkungsquerschnitt von prompten Photonen mit einer transversalen Energie im Bereich von $6 < E_T^\gamma < 15\,\mathrm{GeV}$ und der Pseudorapidität $-1.0 < \eta^\gamma < 2.4$ für Ereignisse mit negativem Viererimpusübertrag $Q^2 < 1\,\mathrm{GeV}^2$ und Inelastizität im Bereich $0.1 < y < 0.7$ werden als Funktion von η^γund E_T^γgemessen.

Zusätzlich werden Wirkungsquerschnitte für Ereignisse mit einem Photon und einem zusätzlichen hadronischen Jet als Funktion der transversalen Energie und Pseudorapidität des Jets und des Impulsbruchteils, der von den wechselwirkenden Partonen getragen wird, gemessen. Die Ereignisse mit einem Photon und hadronischen Jet werden ausserdem benutzt, um die transversalen Korrelationen zwischen dem Jet und dem Photon zu untersuchen. Die Resultate werden mit Vorhersagen von zwei verschiedenen QCD Rechnungen verglichen, welche beide nicht in der Lage sind, die Verteilungen in allen Details zu beschreiben.

Acknowledgments

It is a pleasure to express my gratitude to everyone who supported me during my PhD study time.

First of all, I would like to thank Prof. Dr. Ulrich Straumann who made this whole PhD adventure possible. My time at the Zürich University constitutes the best part of my life. Thank you, I am honoured to be a part of your group, at least for a short time.

Very special thanks go to the person who constantly believed in me, Dr. Katharina Müller. I'll never be able to repay all the kindness and support I received, nor even list all the things I should thank you for. I started my PhD as a self-sufficient person, but I finish deeply indebted. It was a pleasure to know you, Katharina. Thank you for everything.

Deep gratitude makes me to mention my fellow prompt photon student, Carsten Schmitz, who was always ready to share his experience with a confused younger mate. Just as Linus Lindfeld, I'm grateful for the initial guidance in attending the CIP. The detector part of my PhD, though brief, was definitely among the most interesting experiences. Carsten and Linus, Office 222e has never been the same without you two.

Many thanks go to Dr. Stefan Schmitt, the man whose intelligence is only exceeded by the kindness. This is a privilege just to know you, and you gave me much more. I am deeply thankful for all your help, especially with the unfolding procedure. I need to thank all the HaQ working group members, especially the convenors: Dr. Günter Grindhammer, Dr. Daniel Traynor and Dr. Albert Knutsson for all the input to my work. Concerning the publication procedure related to this work, I am obliged to the publication referees, Prof. Dr. Jacek Turnau and Dr. Gerhard Brandt, for all the critical comments they made.

To my dear friend, Deniz Sunar, there is more to say than I could express. Thank you for all the good time we shared. May Allah shine His face upon you, wherever you are. Shiraz Habib, our controversial philosophical discussions were definitively highlihghts of my study, thank you for all those "serious" hours. Guillaume Leibenguth, I knew it would have been easier to make a PhD in physics than to learn how to correctly spell your name, thank you for your friendship. And to Michel Sauter, big thank you for all your positive attitude in the office and fun outside. It is a pleasure to know you. Guys, DESY wouldn't be the same without you all.

I want to show my appreciation to all my friends not directly involved with my PhD, but who contributed by keeping me sane throughout all those years. To my parents, for their deep, unsheakable love and care for me; words can not express my gratitude. Dla moich rodziców, za niezmienna miłość i opieke; slowa nie opisza tego co czuje. Suxi, thank you for being my best friend. Eva, John, Uta, Bona et. al., thank you for being always there for me. Know for sure that you profoundly impacted my life. To Craig Rees, you probably have no idea how much I appreciate that we met. And to all the IBC friends, you all made my days in Hamburg simply unforgettable.

Last, though not least, thanks be to the one who always finds time for me, whenever I need it. To the one who brought me to the place I'm in today. Whatever I say would not have been enough. This is all through you and for you, Jesus. Thank you.

Contents

Introduction

The Standard Model is a well established theory proved to be successful in describing fundamental particles and their interactions. It reduces almost all the variety of subatomic effects into interaction of twelve fundamental particles, six *quarks* and six *leptons*. Three forces, electromagnetic, weak and strong are carried by particles called *bosons*. All the masses are acquired by a coupling to the Higgs boson, the only still undiscovered particle of the Standard Model. Even though the Standard Model is not considered as a final and complete theory, it appears extremely difficult to obtain experimental data contradicting its predictions. It is believed though that *new physics* data needs to be observed at higher energies, as the Standard Model can not for example describe successfully the gravitational force, explain the dark matter / energy phenomena or the observed matter - antimatter asymmetry.

The electron-proton collider, HERA, has been developed particularly to improve the understanding of the strong interaction described by the *quantum chromodynamics* (QCD) theory. The QCD introduces eight massles bosons, called *gluons*, carriers of the strong force. Its strength is given by the coupling constant α_s, which increases with the distance, the effect giving rise to the concept of quark *confinement*. At high energy scales and short distances the strength is sufficiently small to allow perturbative calculations, where scattering amplitudes are expressed in power series of the coupling constant. At larger distances though, QCD calculations become increasingly difficult and processes such as hadronisation, by which partons emerging from high energy collisions form colour neutral hadrons, require further understanding.

Among a large variety of processes which can be studied at HERA, this work concentrates on the prompt photon production. The term *prompt photons* is commonly used to describe photons produced directly in the *hard interaction*, as opposed to secondary photons coming from the decays of unstable particles. Prompt photon measurements, as opposed to measurement based purely on jets are believed to be less sensitive on the details of the non perturbative hadronisation process and can be used to test the QCD predictions. In the region of low four momentum transfer squared between the electron and the proton, so called *photoproduction*, in addition to the *direct* interaction of the mediating photon, the contribution of *resolved* events, where mediating photon fluctuates into a partonic state, becomes more important. Accordingly, prompt photon measurements in photoproduction are sensitive to the hadronic structure of both the proton and the photon. Furthermore, a good understanding of the Standard Model production of isolated photons is required for new physics searches in future experiments, particularly for a search for the Higgs boson in

the two photon decay channel for relatively small Higgs particle masses ($M_H < 140\,\text{GeV}$).

In this thesis the prompt photon measurement in photoproduction is presented. The study is based on data collected by the H1 detector in the years 2004-2007 with an integrated luminosity of 340 pb^{-1}. The study significantly extends the phase space of the previous prompt photon measurement in photoproduction towards photons produced in the forward region of the detector and lower event inelasticities. The extension enhances the contribution of resolved events and consistently improves the sensitivity to the parton density function of the photon.

The main experimental difficulty are secondary photons produced in the decays of unstable neutral hadrons, such as $\pi^0 \to \gamma\gamma$. The prompt photon signal is extracted with a multivariate analysis based on the shape of the energy deposits in the calorimeter. The high statistics of HERA II data used for this analysis, increased by a factor of three compared to the previous measurement, allows to use a sophisticated unfolding method to separate photons from background and determine the cross sections. The production of prompt photons is studied in the inclusive phase space and in the phase space of photons accompanied by the hadronic jet.

The results of this work have been approved by the H1 collaboration, have been shown on several conferences and are in the process of being published.

The outline of this work is as follows:

- Chapter 1, *Theoretical framework*, lays the basis needed for the understanding of this study. General concepts of particle physics are introduced with a special emphasis on the prompt photon production. A review on recent prompt photon measurement is also given.
- Chapter 2, *Theoretical predictions*, describes the QCD calculations and the Monte Carlo predictions used in this analysis.
- Chapter 3, *The H1 experiment at HERA*, provides an overview of the HERA collider and the H1 detector, with the most important detector compontents for this analysis briefly highlighted.
- Chapter 4, *Event reconstruction and preselection*, gives the basic informations of the data used in this analysis and its reconstruction. The trigger used and the selection of electron-proton interactions is also discussed.
- Chapter 5, *Prompt photon selection*, presents the prompt photon in photoproduction event signature with detailed description of cuts applied to obtain the final selection.
- Chapter 6, *Photon signal extraction*, gives an overview of the method employed to discriminate between prompt photon signal and neutral hadron background. It is based on the cluster shapes combined with a multivariate analysis.
- Chapter 7, *Calibration and tuning*, shortly summarises efforts made to improve the Monte Carlo description of the data, particularly to correct the simulation of the cluster shapes which is essential for the final results.

- Chapter 8, *Cross section building*, describes in details the unfolding method employed for a determination of the results. This chapter uses all the pieces introduced previously, such as data selection, Monte Carlo simulation and signal to background discrimination to produce the final cross sections. The measurement errors are also discussed. Furthermore, it describes the correction applied to the QCD calculations.

- Chapter 9, *Results*, is the place where the final cross sections are presented and discussed in comparison to the theoretical predictions.

Chapter 1

Theoretical framework

This chapter introduces the theoretical basis needed for the understanding of prompt photon production in photoproduction. In sections 1.1 and 1.2 the theoretical introduction to the Standard Model and the electron-proton scattering is given. Prompt photon theory and summary of recent studies is presented in sections 1.3 and 1.4. The chapter is finished with discussion of the prompt photon analysis motivation. Finally, the goals for the described work are defined in section 1.5 setting the foundament for the discussion following in next chapters.

1.1 The standard model

The Standard Model (SM) is the experimentally well-tested theory of the particle physics based on fundamental particles and their interactions. Within this model, fundamental particles can be classified according to three basic types: *quarks*, *leptons* and *gauge bosons* (the carriers of force). The quarks and leptons, which in turn are divided into three generations, are members of a family of particles called *fermions* (particles with half integer spin). The properties of these fundamental fermions are summarized in table 1.1. Experimentally six different types of quarks, known as flavours, have been observed: up (u), down (d), charm (c), strange (s), top (t) and bottom (b) are grouped into three generations. The *left handed* (L) quarks, which have the spin oppositely aligned with the direction of motion, are grouped into the doublets. The *right handed* (R) quarks, where the spin is aligned with the direction of motion, are grouped into singlets.

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b \end{pmatrix}_L, \quad u_R, \quad d_R, \quad c_R, \quad s_R, \quad t_R, \quad b_R \tag{1.1}$$

All visible matter is composed of the first generation particles, e.g. protons and neutrons are made out of u and d quarks. Quarks have never been observed as free isolated states, but only exist in bound states, e.g. baryons represent the bound states of three quarks (qqq), and mesons, the bound state of the quark-antiquark ($q\bar{q}$) pairs.

Leptons are spin 1/2 particles which can be observed as free particles. Analogous to the

Gene-	**Quarks**			**Leptons**		
ration	**Flavour**	**Q**	**M [GeV]**	**Flavour**	**Q**	**M [GeV]**
1^{st}	u (up)	$+2/3$	$2.55^{+0.75}_{-1.05} \times 10^{-3}$	e (electron)	-1	5.1×10^{-4}
	d (down)	$-1/3$	$5.04^{+0.96}_{-1.54} \times 10^{-3}$	ν_e (e-neutrino)	0	$< 1 \times 10^{-8}$
2^{nd}	c (charm)	$+2/3$	$1.27^{+0.07}_{-0.11} \times 10^{0}$	μ (muon)	-1	0.105
	s (strange)	$-1/3$	$104^{+26}_{-34} \times 10^{-3}$	ν_μ (μ-neutrino)	0	$< 2 \times 10^{-4}$
3^{rd}	t (top)	$+2/3$	$171.2^{+2.1}_{-2.1} \times 10^{0}$	τ (tau)	-1	1.776
	b (bottom)	$-1/3$	$4.20^{+0.17}_{-0.07} \times 10^{0}$	ν_τ (τ-neutrino)	0	$< 2 \times 10^{-2}$

Table 1.1: The properties of the fundamental fermions (quarks and leptons, spin = 1/2) of the SM [1, 2]. The anti-particle partners of these fermions (not included in the table) have the same mass (M), but opposite electric charge (Q). Q is given in units of the proton charge.

quarks, there are 3 known generations which differ from each other only in mass and flavour. The electron (e), muon (μ) and tau (τ) particles each have an associated low mass, chargeless neutrino. Leptons are also group into singlets and doublets:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L, \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L, \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L, \quad e_R, \quad \mu_R, \quad \tau_R \tag{1.2}$$

The neutrinos, neutral leptons, are considered to be massless within the SM. The observation of neutrino mixing [3–7] however, has shown that in fact they can not be massles. The electron, like the proton, is a stable particle and is present in almost all matter. The μ and τ particles are unstable and are found primarily in cosmic rays.

Important ingredients of the SM are the intermediate gauge bosons, or the carriers of force. Table 1.2 lists the fundamental forces and their carriers. The gauge bosons transmit three of the four fundamental forces through which matter interacts. The gluon (g) is responsible for the strong force, which binds together quarks inside protons and neutrons, and holds together protons and neutrons inside atomic nuclei. The photon (γ) is the electromagnetic force carrier that governs electron orbits and chemical processes. The photon couples to the electric charge. Lastly, the weak force is mediated by $W^\pm$ and Z^0 bosons responsible for radioactive decays. The weak force couples to quarks as well as to leptons. Due to the lack of colour and electric charge, neutrinos do not interact via the strong or the electromagnetic force, and therefore interact with matter only via the weak interactions. Within the SM, theories of electromagnetic and of weak interactions are unified to the Electroweak theory by Glashow, Salam and Weinberg [8].

The SM includes the strong, electromagnetic and weak forces and all their carrier particles, and describes how these forces act on all the matter particles. However, the fourth force,

Force	Force carriers	Coupling	Relative strength	Mass [GeV]	Spin
Strong	g (gluon)	quarks	1	0	1
Electromagnetic	γ (photon)	charged particles	1.4×10^{-2}	$< 3 \times 10^{36}$	1
Weak	$W^{\pm}$, Z^0	quarks and leptons	2.2×10^{-6}	80, 91	1
Gravitational	graviton	all particles	10^{-38}	0	2

Table 1.2: The fundamental forces and force carriers, i.e gauge bosons [1]. Three kinds of fundamental forces are combined in the SM. The fourth fundamental interaction, gravity, is shown separately as it is not yet included in SM.

gravity, is not part of the SM. In fact, attempts have been made to describe gravity as a quantum field theory, with the interaction mediated by a spin-2 boson, the *graviton*, associated to the gravitational field. Such extensions of the SM, however, need also an adapted description of the SM itself, because of problems such as non-renormalizability. The gravitational force does not play a significant role in atomic and subatomic processes because it is much weaker than the SM forces.

Although the SM explains the interactions among quarks, leptons, and bosons, the theory does not include an important property of elementary particles, their mass. In 1964 Francois Englert and Robert Brout [9], Peter W. Higgs [10], and independently Gerald Guralnik, C. R. Hagen, and Tom Kibble [11], proposed a mechanism that provides a way to explain how the fundamental particles could acquire mass. The theory states that the space is filled with a scalar field, now called the Higgs field. Particles moving through the field acquire mass. The particle associated with the Higgs field is the Higgs boson, having no intrinsic spin or electric charge. The Higgs boson does not mediate a force as do the other gauge bosons and it has not been observed yet. Finding it, is the key to discover whether the Higgs field exists, and whether the Higgs mechanism for the origin of mass is indeed correct. Detectors at Fermilab and eventually at the LHC[1] at CERN [2] are looking for the elusive Higgs particle.

1.2 Basics of electron - proton scattering

Electroweak theory describes electron - proton scattering by the exchange of the virtual gauge bosons. *Neutral current* (NC) interactions are mediated by either a photon γ or neutral boson Z^0 in opposition to *charged current* (CC) interactions with exchanged

[1] **L**arge **H**adron **C**ollider.
[2] **C**onseil **E**uropéen pour la **R**echerche **N**ucléaire.

a) b)

Figure 1.1: Lowest order Feymann diagrams for the neutral current (a) and the charged current (b) electron - proton scattering process.

For a fixed centre of mass energy $\sqrt{s}$ the kinematics of the reaction is fully given by the four-momenta of the scattered particles. If k and k' denote the four-momenta of the scattered lepton, P corresponds to the four-momentum of the incoming proton and both electron and proton masses can be neglected, s can be expressed by

$$s = (k + P)^2 = 4E_e E_p, \tag{1.3}$$

where E_e and E_p denote the energy of the electron and proton beams.

The four-momentum transfer $q = k - k'$ defines the invariant mass of the virtual boson $\sqrt{q^2}$. Since q^2 has a negative value, the commonly used variable is a negative four-momentum transfer squared

$$Q^2 = -q^2 = -(k - k')^2. \tag{1.4}$$

The Q^2 is also referred to as a *virtuality* of the exchanged boson and in case of $Q^2 \approx 0$ the exchanged boson is a quasi-real or on mass-shell particle.

The relative energy transfer of the electron in the proton rest frame is described by the *inelasticity* y defined as

$$y = \frac{q \cdot P}{k \cdot P}. \tag{1.5}$$

The Björken scaling variable x is defined as

$$x = \frac{Q^2}{2P \cdot q}. \tag{1.6}$$

If the struck parton is colinear to the proton, the Björken scaling variable x represents the fraction of the four-momentum of the proton carried by the struck quark. Both x and y take values between zero and one.

Neglecting particle masses, the quantities given above may be related by

$$Q^2 = x \cdot y \cdot s. \tag{1.7}$$

Figure 1.2 shows the neutral and charged current cross sections measured by the H1 experiment as a function of Q^2. Below $Q^2 \approx 3000\,\mathrm{GeV}$ the neutral current cross section clearly dominates. Due to high masses of the Z^0 and $W^\pm$ bosons, at lower Q^2 their exchanges are kinematically suppressed. In the following only the pure photon exchange is considered. The differential NC cross section for the process $e^\pm p \to e^\pm X$ is given by

$$\frac{\mathrm{d}^2\sigma^\pm_{NC}}{\mathrm{d}x\mathrm{d}Q^2} = \frac{2\pi\alpha^2}{xQ^4} Y_+ \left(F_2(x,Q^2) - \frac{y^2}{Y_+} F_L(x,Q^2) \right). \tag{1.8}$$

Here, α denotes the fine structure constant, $Y_+ = (1 + (1-y)^2)$ is the *helicity factor* and $F_2(x,Q^2)$ and $F_L(x,Q62)$ are called proton structure functions, which parametrise the proton content as probed by the virtual photon. The F_L contribution is kinematically suppressed compared to F_2 and becomes significant only at very high event inelasticities y.

1.2.1 Quark - parton model

In the infinite momentum frame ($P^2 \gg m_P^2$), the proton can be considered as a parallel stream of independent partons, from which every one carry the fraction $\xi_{p,i}$ of the longitudinal proton momentum. This picture is used in the quark parton model (QPM). Deep inelastic scattering processes can then be interpreted as elastic electron scattering on a single parton, as visualised in the figure 1.3. Other partons, not participating in the hard interaction form the *proton remnant* and are referred to as *spectator partons*. The individual partons within the proton are not directly visible and the proton content can be described by universal probabilistic parton densities. Since the partons in the proton have been identified as quarks and gluons, for each quark flavour and gluon parton density functions (PDFs) exist, which give the probability of finding a parton i with a momentum fraction ξ_p in the proton.

In the QPM, F_2 depends only on x and can be written as

$$F_2(x) = x \sum_q e_q^2 \left[f_q^p(x) + f_{\bar{q}}^p(x) \right], \tag{1.9}$$

where the sum runs over all the quark flavours q, e_q are the quark charges and the functions f_q^p and $f_{\bar{q}}^p$ contain the quark and antiquark densities in the proton. In this picture the longitudinal structure function F_L disappears

$$F_L(x) = 0. \tag{1.10}$$

Early experimental results on F_2 were effectively in agreement with the so-called *scaling* behaviour of F_2 (no Q^2 dependence of F_2). Later, from the observation of the *scaling violation* at lower x values, it was concluded that also gluons and gluon splitting into quark-antiquark pairs have to be considered for the successful description of the proton content.

Figure 1.3: Electron - proton scattering in the quark parton model.

Figure 1.4: Higher order loop corrections to the scattering process. Gluon loop (a) and fermion loop (b).

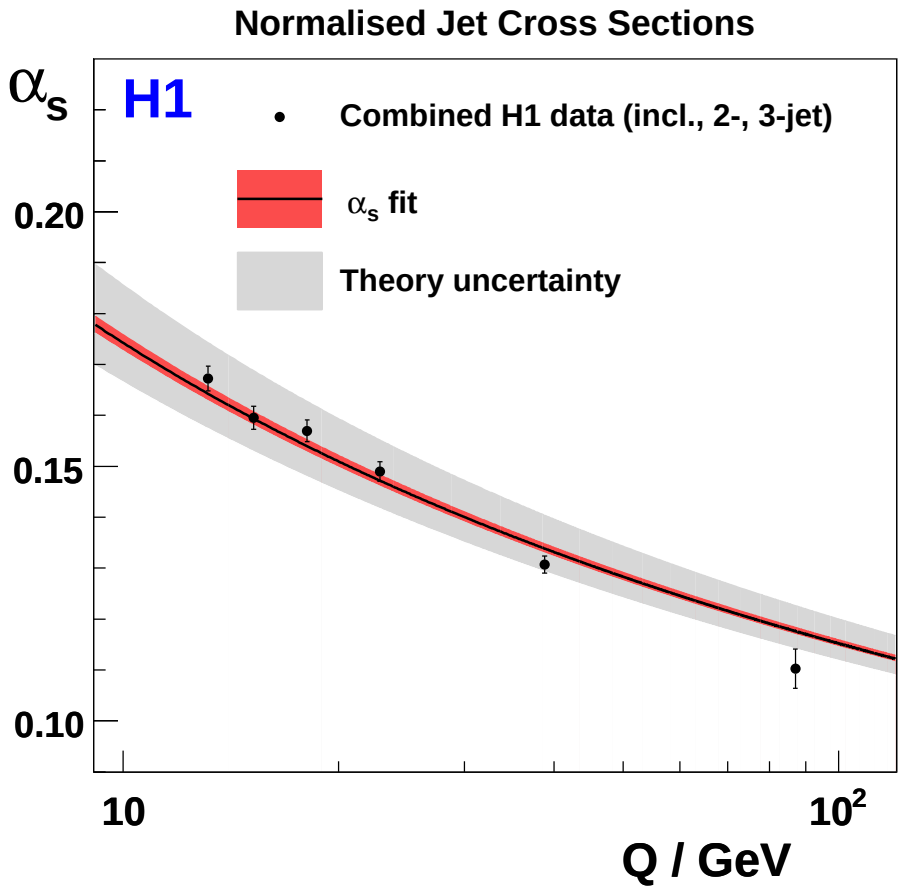

Figure 1.5: The scale dependence of the strong coupling constant α_s. Figure taken from [13].

1.2.3 Factorisation theorem

The calculation of gluon radiations ($q \to qg$, $g \to gg$) and splittings ($g \to q\bar{q}$) in pQCD gives rise to infrared divergencies. The introduction of a factorisation scale μ_F cures the problem in a similar way as the μ_R does. Above the cut-off scale μ_F collinear soft radiations are absorbed into the parton density functions. Reliable pQCD calculations are possible only for $\mu_F \gg \Lambda_{QCD}$.

Within the *factorisation theorem*, the proton structure function F_2 can be written as a convolution of perturbatively calculable coefficient functions C^i and parton density functions $f_{i/p}(\xi)$, which could be interpreted as a probability of finding a parton of type i carrying a fraction ξ of the protons longitudinal momentum:

$$F_2(x,Q^2) = x \sum_{i=q,\bar{q},g} \int_x^1 \mathrm{d}\xi C^i \left(\frac{x}{\xi}, \frac{Q^2}{\mu_R^2}, \frac{\mu_F^2}{\mu_R^2}, \alpha_s(\mu_R^2) \right) f_{i/p}(\xi, \mu_F^2, \mu_R^2). \tag{1.12}$$

Based on the factorisation theorem, the *ep* cross section can be decomposed in the presence

of a hard scale, into

$$\sigma_{ep} \sim f_{i/p}(\xi, \mu_F^2) \otimes \hat{\sigma}_{i,j}(\hat{s}, \alpha_s(\mu_R^2), \mu_R, \mu_F) \otimes f_{i/\gamma}(\xi_\gamma, \mu_F^2). \tag{1.13}$$

where $\hat{\sigma}$ denotes a hard subprocess calculable in pQCD and $f_{i/p}$ and $f_{i/\gamma}$ are parton density functions of proton and photon.

1.2.4 Parton evolution model

A calculation of $f_{i/p}$ is beyond the scope of perturbative QCD and it should be extracted from experimental measurements. Nevertheless, the dependence of $f_{i/p}$ can be studied theoretically, given that the cross sections must not depend on μ_F. This leads to *parton evolution equations* which are used to evolve parton density functions known at a certain starting scale μ_0 up to the factorisation scale μ_F. The scale μ_0 is usually taken as the momentum transfer in the hard interaction, $\mu_0 = Q^2$. In order to solve parton evolution equations, some approximations must be made, which are expected to be valid only in certain regions of the phase space. Parton evolution equations can be obtained by gluon splitting and gluon radiation processes leading to a *gluon ladder*. The ladder diagram of gluon emissions is shown in figure 1.6. In the following, three main approximations are described.

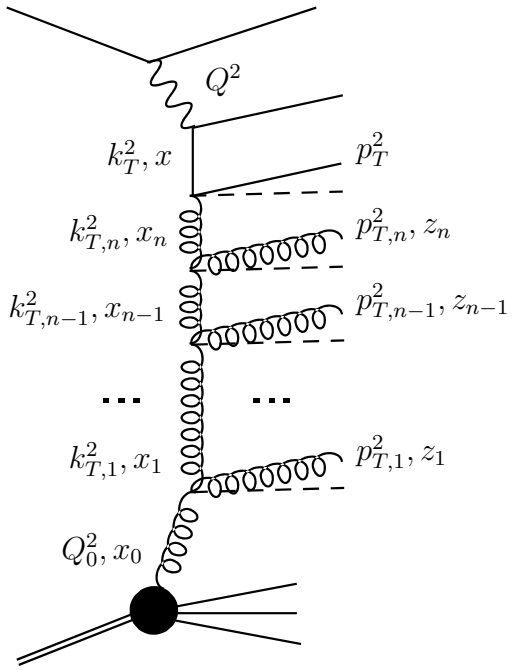

Figure 1.6: A schematic gluon ladder diagram of parton evolution in *ep* scattering. A parton from the proton interacts with a virtual photon from the electron after radiating n gluons. The transverse and longitudinal momenta of each emitted gluon is denoted $p_{T,i}$ and z_i, respectively.

1.2.4.1 DGLAP evolution equations

The Dockshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [14–18] define the way in which quark and gluon momentum distribution evolve with the scale of the interaction. Within DGLAP approach a strong ordering of the transverse momenta $k_{T,i}$

$$Q^2 \gg k_{T,i}^2 \gg k_{T_i-1}^2 \gg k_{T,i-2} \ldots \gg k_{T_1} \gg Q_0^2, \tag{1.14}$$

and a soft ordering of the fractional longitudinal momenta x_i

$$x_i < x_{i-1} < x_{i-2} \ldots < x_1 \tag{1.15}$$

are assumed. Here, Q_0^2 is the virtuality of the parton at the start of the emission cascade and Q^2 is the virtuality of the exchanged photon. The DGLAP evolution equations are given by

$$\begin{aligned} &\frac{\partial f_{q/p}(x,Q^2)}{\partial \log Q^2} = \\ &= \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{\mathrm{d}y}{y} \left[f_{q/p}(y,Q^2) P_{qq}\left(\frac{x}{y}\right) + f_{g/p}(y,Q^2) P_{qg}\left(\frac{x}{y}\right) \right], \end{aligned} \tag{1.16}$$

$$\begin{aligned} &\frac{\partial f_{g/p}(x,Q^2)}{\partial \log Q^2} = \\ &= \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{\mathrm{d}y}{y} \left[f_{q/p}(y,Q^2) P_{gq}\left(\frac{x}{y}\right) + f_{g/p}(y,Q^2) P_{gg}\left(\frac{x}{y}\right) \right], \end{aligned} \tag{1.17}$$

where q and g denote quark and gluon density function respectively and P_{ij} are *splitting functions* of a parton i to parton j with the momentum fraction $\left(\frac{x}{y}\right)$ shown in figure 1.7. Both DGLAP equations assume massles partons and are hence only valid for gluons and light quarks (u, d and s). The splitting functions are given in leading order by

$$P_{qq}(z) = \frac{4}{3}\left(\frac{1+z^2}{(1-z)_+}\right) + 2\delta(1-z), \tag{1.18}$$

$$P_{qg}(z) = \frac{1}{2}\left(z^2 + (1-z)^2\right), \tag{1.19}$$

$$P_{gq}(z) = \frac{4}{3}\left(\frac{1+(1-z)^2}{z}\right), \tag{1.20}$$

$$P_{gg}(z) = 6\left(\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right) + \left(\frac{11}{2} - \frac{n_f}{3}\right)\delta(1-z), \tag{1.21}$$

where the notation $(F(x))_+$ defines a distribution such that for any sufficiently regular test function $f(x)$,

$$\int_0^1 dx f(x) \left(F(x)\right)_+ = \int_0^1 dx \left(f(x) - f(1)\right) F(x). \tag{1.22}$$

Figure 1.7: Feymann diagrams for four splitting functions in the DGLAP evolution equations.

1.2.4.2 BFKL evolution equations

The DGLAP equations neglect terms of the form $\log(1/x)$ which may become large as x becomes small. Summation of such contributions leads to unintegrated gluon distributions (dependent on the transverse momentum k_T), which obey the Balitsky-Fadin-Kuraev-Lipatov (BFKL) [19, 20] approximation. In the framework of the unintegrated gluon distribution, predictions for the measured cross sections are calculated using the k_T-factorisation theorem [21]. Cross sections are factorised into an off-shell (k_T dependent) partonic cross section and a k_T-unintegrated parton distribution. The BFKL approximation allows the summation of terms with leading powers of $\alpha_s \log\frac{1}{x}$ in the regime of very low x and moderate Q^2. In this approach, a strong ordering of the fractional longitudinal momentum x_i

$$x_i \ll x_{i-1} \ll x_{i-2} \ll ... \ll x_1 \tag{1.23}$$

and no ordering on the transverse momenta k_T along the ladder are assumed. The resulting BFKL evolution equation is given by

$$\frac{\partial f_{g/p}(x,k_T^2)}{\partial \log(1/x)} = \frac{3\alpha_s}{\pi} k_T^2 \int_0^\infty \frac{\mathrm{d}k_T'^2}{k_T'^2} \left[\frac{f_{g/p}(x,k_T'^2) - f_{g/p}(x,k_T^2)}{|k_T'^2 - k_T^2|} + \frac{f_{g/p}(x,k_T^2)}{\sqrt{4k_T'^4 + k_T^4}} \right]. \tag{1.24}$$

The terms within brackets of equation 1.24 correspond to the real gluon emission and virtual corrections, respectively. The BFKL gives the evolution of $f_{g/p}(x,k_T^2)$ with respect to small x. It can be solved for any small x and k_T once it is known for a starting value. For a fixed α_s, the BFKL equation can be solved and the result is

$$f(x,k_T^2) = F(x,k_T^2) \left(\frac{x}{x_0}\right)^{-\lambda}, \tag{1.25}$$

where $\lambda = \frac{3\alpha_s}{\pi} 4\log 2$. Hence, this approximation predicts that the gluon density increases proportional to $(1/x)^{-\lambda}$ as x decreases.

1.2.4.3 CCFM evolution equations

Both the DGLAP and the BFKL methods only sum over one particular leading behaviour of the evolution to obtain results. A complete (infinite order) calculation should take into account both, the terms in $\log(Q^2)$ and $\log(1/x)$ and sum over them. To accomplish this, Ciafaloni [22] and Catani, Fiorani and Marchesini [23] introduced angular ordering

for the emitted gluons. The maximum allowed angle is defined by the hard scattering, where the quark pair is produced. This is combined with unintegrated gluon densities and off-shell partons, like in the BFKL approach. This method seems very promising, as it can (approximately) reproduce the DGLAP and BFKL equations within the appropriate limits. However, it is still incomplete. Currently, it does not include quark evolution contribution.

1.2.5 Proton structure

The parton density functions of the proton are not physical observables and need to be determined from fits to the measured $F_2(x, Q^2)$ data. The quark densities are very well known in the kinematical range accessible by the present experiments, while the gluon density has still uncertainties of the order of 10% to 40%. In figure 1.8 the parton density of the proton as a function of x for a fixed $Q^2 = 1.9\,\mathrm{GeV}^2$ and $Q^2 = 10\,\mathrm{GeV}^2$ is shown, as extracted with the help of recent H1 data [24]. It can be seen that the valence quarks u and d dominate at high x, whereas the gluons come into play at lower x exceeding the quark contribution by far at low enough x. This behaviour is correlated with the large scaling violations of $F_2(x, Q^2)$ observed at low x. In the linear scale plots the gluon contribution is scaled down by factor 20.

Figure 1.9 shows the proton structure F_2 as a function of Q^2 for various x values as obtained by the H1 collaboration [25]. Clear scaling violation can be observed. The radiation of a gluon reduces the original momentum fraction of the scattering quark and in addition gluons can split into quark-antiquark pairs with relatively small momentum fractions of the proton. At higher momentum transfers more such processes can be resolved and hence the quark densities are expected to rise with Q^2 at low x and to decrease with Q^2 at high x.

A lot of PDF sets from different groups are available which are based on global fits to the data, primary inclusive DIS data. In this analysis the results of two different proton PDF groups are used, the parton distribution functions fitted by CTEQ [26] group and by MRST [27] group.

1.2.6 Photoproduction

In the region of small Q^2 electron-proton scattering can be simplified by factorising radiation of a photon from the electron and the subsequent interaction of the photon with the proton. This particular domain of electron-proton scattering is called *photoproduction*. Since the photon propagator gives rise to a factor $1/Q^2$ in the inclusive NC DIS cross section, photoproduction dominates the *ep* scattering cross section. In photoproduction processes, the electron radiates a quasi-real ($Q^2 < 1\,\mathrm{GeV}^2$) photons with energy fractions y according to the Weizsäcker-Williams approximation. The variable y is directly related to the centre of mass energy $W_{\gamma p}$ of the photon proton system:

$$W_{\gamma p}^2 = (q + P)^2 = y \cdot s - Q^2 \approx y \cdot s, \tag{1.26}$$

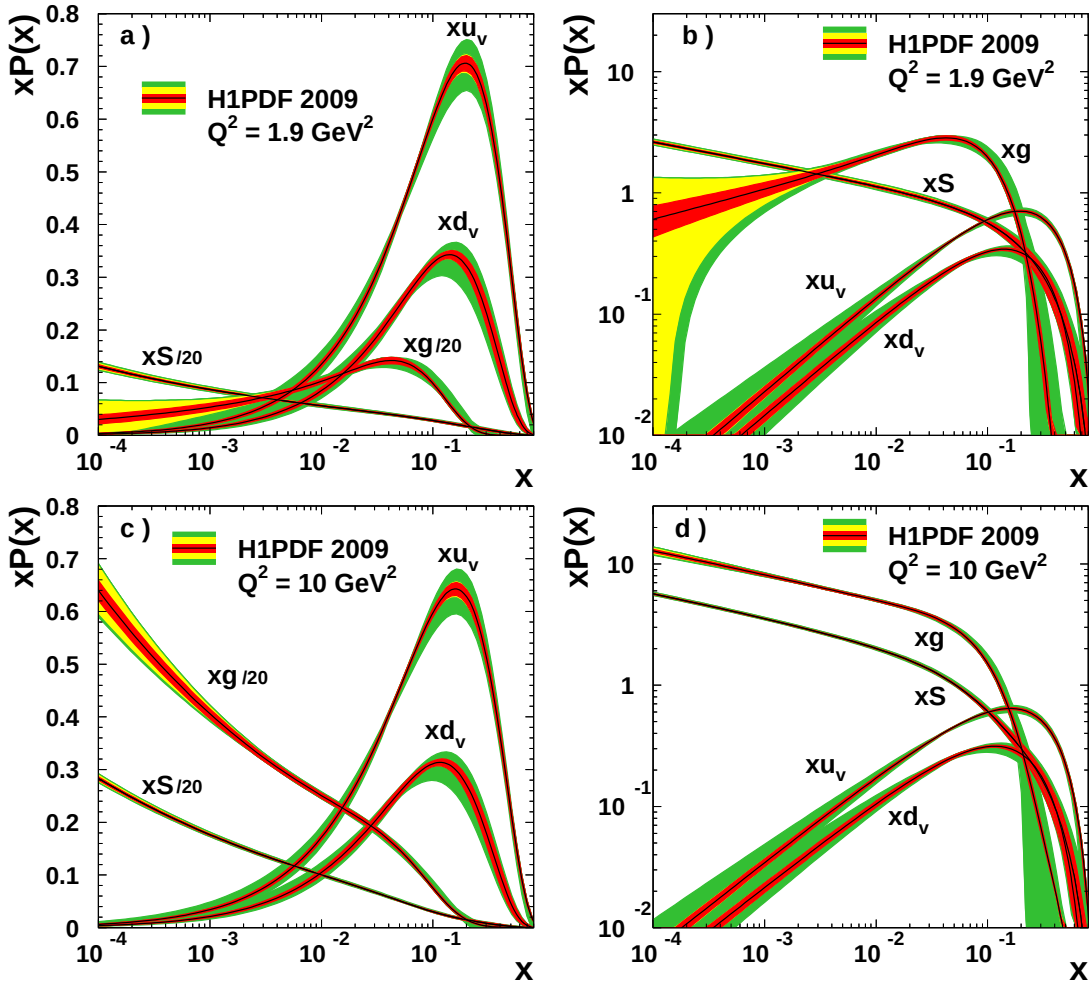

Figure 1.8: Parton distributions as determined by the H1PDF 2009 QCD fit at $Q^2 = 1.9\,\mathrm{GeV}^2$ (a, b) and at $Q^2 = 10\,\mathrm{GeV}^2$ (c, d). In a) and c) (linear vertical scale), the gluon and sea-quark densities are downscaled by a factor 0.05. The inner error bands show the experimental uncertainty, the middle error bands include the theoretical model uncertainties of the fit assumptions, and the outer error band represents the total uncertainty including the parameterisation uncertainty. The figure is taken from [24].

where P is the proton and q the photon four-momentum. The proton and electron masses are here and in the following neglected.

The cross section $\sigma_{ep\to eX}$ can be factorised into two parts:

$$\sigma_{ep\to eX} = \int \mathrm{d}y f_{\gamma/e}(y)\sigma_{ep\to eX}(y), \tag{1.27}$$

where $f_{\gamma/e}$ is the *photon flux*. Using the Weizsäcker-Williams approximation [28,29], which neglects terms involving longitudinal photon polarisation, the flux of photons with energy fraction y and up to virtuality Q^2_{max} can be calculated via

$$f_{\gamma/e}(y) = \frac{\alpha}{2\pi}\left[\frac{1+(1-y)^2}{y}\ln\left(\frac{Q^2_{max}}{Q^2_{min}}\right) - 2m_e^2 y\left(\frac{1}{Q^2_{min}} - \frac{1}{Q^2_{max}}\right)\right]. \tag{1.28}$$

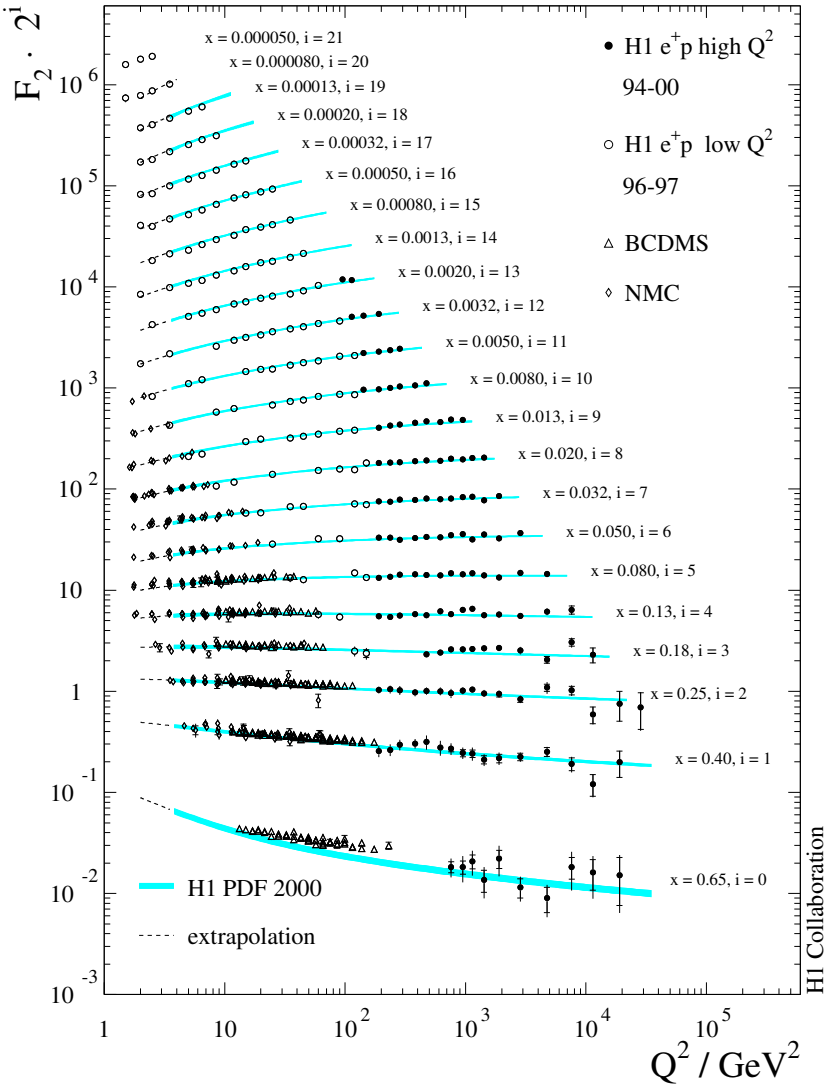

Figure 1.9: The proton structure function F_2 as a function of Q^2 for different values of x measured by the H1 collaboration and the fixed target experiments BCDMS and NMC. The results are compared with the corresponding Standard Model expectation determined from the H1 PDF 2000 fit indicated by the error bands. The figure is taken from [25].

The fine structure constant is denoted as α and $Q^2_{min} = \frac{m_e^2 y}{1-y}$ is the kinematic lower limit. The maximum photon virtuality Q^2_{max} is given by the scattered electron energy E'_e and angle θ_e:

$$Q^2_{max} = 2E_e E'_e(1 + \cos\theta_e), \tag{1.29}$$

with the energy of the incoming electron denoted as E_e.

Figure 1.10: Direct (a) and resolved (b) photoproduction event. The hard interaction is described by the matrix element (ME) and the proton and photon content by the parton density function (pdf).

The momentum fraction x_γ is defined as

$$x_\gamma = \frac{P \cdot p_j}{P \cdot q}, \tag{1.31}$$

where p_j is the four-momentum of the parton inside the photon. Summing over all colours

Figure 1.11: Quark density of the photon, figure taken from [30].

The QCD corrections have been studied [31]. They preserve the logarithmic dependence on μ_γ^2 and thus still the equation 1.30 gives qualitatively valid description of the difference between various quark contributions. Photon quark densities have been measured mainly in e^+e^- collisions. A nice review can be found in [32]. The photon structure function F_2^γ has been investigated up to large scales, but the experimental uncertainties at high x_γ and high scales are still large. The knowledge about photon gluon density is poorer, but the observed F_2^γ scaling violations, compared to the same effect for the proton F_2 suggest that the gluon density in photon is much less pronounced and the resolved photon consist mainly of quarks.

1.3 Prompt photon production

Prompt photon measurements complement the measurements relying on jets. Complicated, non perturbative hadronisation process affects prompt photon measurements to a

lower extend compared to jets. Also, the energy resolution is usually higher for photon measurements. Compared to the jet production, prompt photon processes are suppressed by order α due to the coupling of the photon. Therefore, the cross section is much lower and a measurement faces difficulties with high rates of background photons. As a result, complicated background discriminator methods are required and may lead to high systematic error.

Figure 1.12 shows some leading order Feymann diagrams of prompt photon production in both direct and resolved processes, while some higher order Feyman diagrams are shown in figure 1.13, particularly a box diagram (figure 1.13a) and the photon radiated directly from the quark line (figure 1.13b).

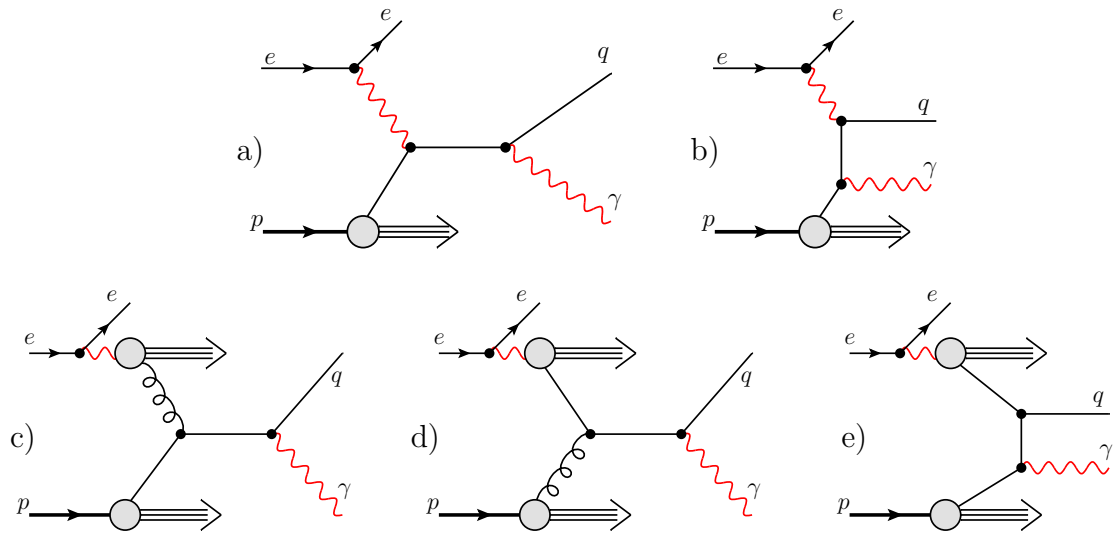

Figure 1.12: Direct (a, b) and resolved (c - e) prompt photon process in leading order.

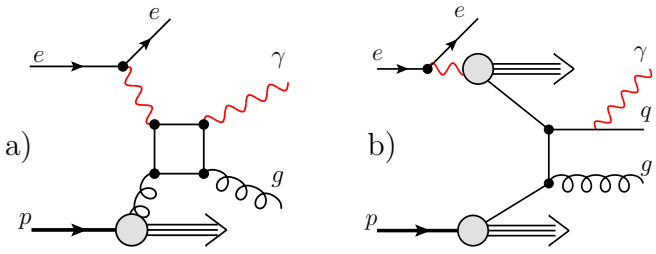

Figure 1.13: Box diagram for the prompt photon production (a) and the photon radiated from the quark line (b).

In addition to photons coming directly from the hard interaction, photons in the final state can originate from the quark-to-photon fragmentation process. The examples are depicted in figure 1.14.

Prompt photon production is sensitive to the parton structure of the proton and due to resolved processes, to the parton content of the photon.

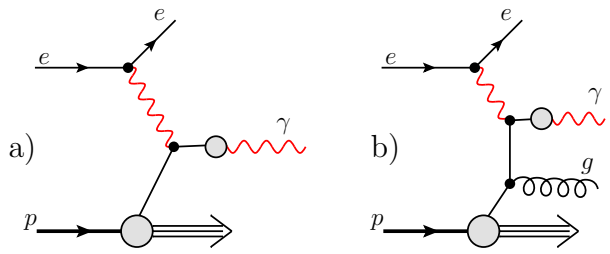

Figure 1.14: Examples for the fragmentation process at LO (a) and NLO (b).

Following the distinction above, the prompt photon cross section can be divided into four parts: direct non-fragmentation $\sigma_{dir}^{nonfrag}$, direct fragmentation σ_{dir}^{frag}, resolved non-fragmentation $\sigma_{res}^{nonfrag}$ and resolved fragmentation σ_{res}^{frag} and factorised in the following way:

$$\mathrm{d}\sigma_{dir}^{nonfrag} = \sum f_{a/p}(x,\mu^2) \otimes f_{\gamma/e}(y) \otimes \sigma_{a\gamma\to\gamma X}, \tag{1.33}$$

$$\mathrm{d}\sigma_{dir}^{frag} = \sum f_{a/p}(x,\mu^2) \otimes f_{\gamma/e}(y) \otimes \sigma_{a\gamma\to dX} \otimes D_{\gamma/d}(z), \tag{1.34}$$

$$\mathrm{d}\sigma_{res}^{nonfrag} = \sum f_{a/p}(x,\mu^2) \otimes f_{\gamma/e}(y) \otimes f_{b/\gamma}(x_\gamma,\mu^2) \otimes \sigma_{ab\to\gamma X}, \tag{1.35}$$

$$\mathrm{d}\sigma_{res}^{frag} = \sum f_{a/p}(x,\mu^2) \otimes f_{\gamma/e}(y) \otimes f_{b/\gamma}(x_\gamma,\mu^2) \otimes \sigma_{a\gamma\to dX} \otimes D_{\gamma/d}(z). \tag{1.36}$$

Here, $f_{a/p}(x,\mu^2)$ and $f_{b/\gamma}(x_\gamma,\mu^2)$ is the already introduced probability of finding parton a in a proton and a photon respectively given longitudinal momentum fraction x and a scale probing the partonic structure μ. Similarly, $f_{\gamma/e}(y)$ is the photon flux calculated with the Weiszäcker-Williams approximation (see equation 1.28). $D_{\gamma/d}(z)$ describes the fragmentation of the parton into the photon.

The parton fragmentation function $D_{\gamma/d}(z)$ gives the probability that the parton d will produce a photon γ carrying a fraction z of the parton momentum. It should be noted that already at leading order a collinear singularity appears in the photon emission by the quark. As physical cross sections are necessarily finite, the singularity may be factorised into the fragmentation function defined at the factorisation scale $\mu_{F,\gamma}$. Within the so-called phase space slicing method [33] a parameter y_{min} can be introduced, which separates the divergent collinear contribution from the finite contribution, where the outgoing quark and photon are still theoretically resolved. In this context the quark-to-photon fragmentation function at order α is given by [34]:

$$D_{\gamma/d}(z) = D_{\gamma/d}(z,\mu_{F,\gamma}) + \frac{\alpha e_q^2}{2\pi}\left(P_{q\gamma}^{(0)}(z)\ln\frac{z(1-z)y_{min}s_{eq}}{\mu_{F,\gamma}^2} + z\right), \tag{1.37}$$

where $D_{\gamma/d}(z,\mu_{F,\gamma})$ describes the non perturbative transition $q\to\gamma$ at the factorisation scale $\mu_{F,\gamma}$. The second term represents the finite part after absorption of the collinear quark-photon contribution into the bare fragmentation function separated by the parameter y_{min}. The variable e_q denotes the charge of quark q and s_{eq} the electron-quark centre of

mass energy squared. In the second term also the LO quark-to-photon splitting function $P_{q\gamma}^{(0)} = (1 + (1-z)^2)/z$ contributes.

The variation of the fragmentation function $D_{\gamma/d}(z, \mu_{F,\gamma})$ with the scale $\mu_{F,\gamma}$ can be predicted by the evolution equations analogue to the DGLAP evolution equations. The fragmentation function $D_{\gamma/d}(z, \mu_{F,\gamma})$ at the scale $\mu_{F,\gamma}$ can be related to the one at the initial scale μ_0 by

$$D_{\gamma/d}(z, \mu_{F,\gamma}) = \frac{\alpha e_q^2}{2\pi} P_{q\gamma}^{(0)}(z) \ln\left(\frac{\mu_{F,\gamma}^2}{\mu_0^2}\right) + D_{\gamma/d}(z, \mu_0). \tag{1.38}$$

1.4 Recent prompt photon measurements

In this section some recent results on the production of prompt photons are presented. At first some measurements using HERA data, both in deep inelastic scattering and photoproduction are reviewed. A discussion on recent prompt photon production results measured in hadron collisions follows.

1.4.1 Results at HERA

At HERA inclusive prompt photon and photon + jet production was measured both in photoproduction region and in DIS.

In photoproduction both H1 [35] and ZEUS [36–38] have already reported prompt photon measurements, which were compared to PYTHIA and HERWIG Monte Carlo (MC - see section 2.1) as well as to NLO calculations. Figure 1.15 presents the H1 and ZEUS [3] inclusive prompt photon cross sections in bins of transverse energy of the photon E_T^γ and its pseudorapidity [4] η^γ compared to the MC predictions. The kinematical range of the measurements is given by $Q^2 < 1\,\mathrm{GeV}^2$, $0.2 < y < 0.7$, $-1.0 < \eta^\gamma < 0.9$ and $5 < E_T^\gamma < 10\,\mathrm{GeV}$. Slightly different ZEUS measurement phase space has been corrected for. The ZEUS measurement was performed using $38.4\,\mathrm{pb}^{-1}$ of data integrated luminosity, while H1 was based their results on $105\,\mathrm{pb}^{-1}$ of data integrated luminosity. The two measurements are consistent within errors, but predictions are low in normalisation by $40-50\%$. The comparison of the H1 measurement with the QCD calculation is presented in figure 1.16a and 1.16b. After the hadronisation and multiple interaction correction, the shape of the cross section is reasonably well described, but the normalisation is again too low by $30-40\%$.

The prompt photon cross sections with the additional requirement on the accompanying hadronic jet as measured by the H1 collaboration [35] are presented in figure 1.16c and 1.16d. The measurement is somehow better described by the calculation in this case. The better description, together with the fact that NLO corrections are in average smaller

[3] The ZEUS measurement was corrected to the kinematic range used by the H1 analysis.

[4] For a definition of a pseudorapidity the reference system is needed. It is introduced in section 3.2 where the η definition is also given.

than in the inclusive case, suggests that higher order corrections are of less importance for prompt photons if accompanied by an energetic jet.

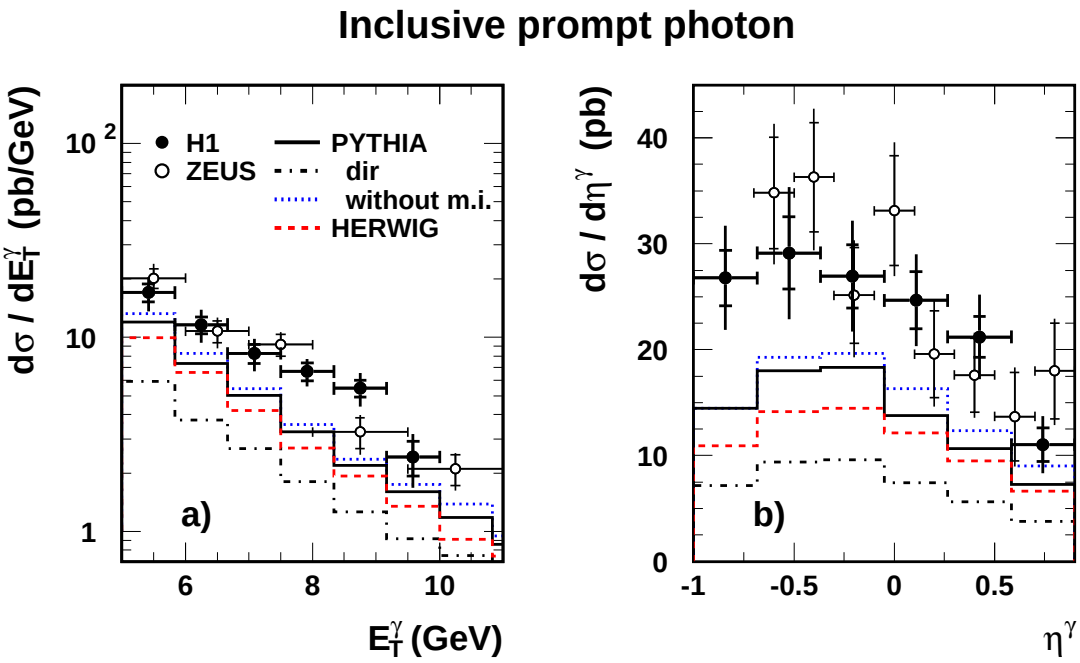

Figure 1.15: Inclusive prompt photon differential cross sections $d\sigma/dE_T^\gamma$ (a) and $d\sigma/d\eta^\gamma$ (b) in photoproduction as measured by the H1 [35] and ZEUS [36] collaborations compared to predictions by PYTHIA (solid black line) and HERWIG (dashed red line). The contribution of direct events is presented with a dashed dotted line and a prediction without multiparton interaction with blue dotted line. Cross sections of the ZEUS were corrected to the kinematics used by the H1 analysis.

Prompt photon production with accompanying jet was also studied by the ZEUS collaboration [38] and some results of this study are presented in figure 1.17. Monte Carlos again predict cross sections low in the normalisation, while QCD calculations give a reasonable description of the data. Especially the calculation based on the k_T factorisation approach (see section 2.2.2) performs well, while NLO calculations predict too low cross sections for low transverse energies of the photon and low pseudorapidities. For an increased minimum transverse energy of the photons to 7 GeV, the differential cross sections are well described by all studied calculations.

In deep inelastic scattering, prompt photons were studied by both ZEUS [41, 42] and H1 [43]. Figure 1.18 presents the differential cross sections measured by the ZEUS collaboration [41] compared to MC predictions of PYTHIA and HERWIG. The measurement has been performed in the phase space $5 < E_T^\gamma < 10\,\text{GeV}$, $-0.7 < \eta^\gamma < 0.9$ and $Q^2 > 35\,\text{GeV}^2$. The transverse energy distribution is reasonably described by PYTHIA scaled by factor 2.3 and by HERWIG scaled by factor 7.9. PYTHIA fails to predict even the shape of cross section in bins of η^γ, while HERWIG predicts a too steep Q^2 slope. The contribution of photons radiated from the electron was neglected in this study.

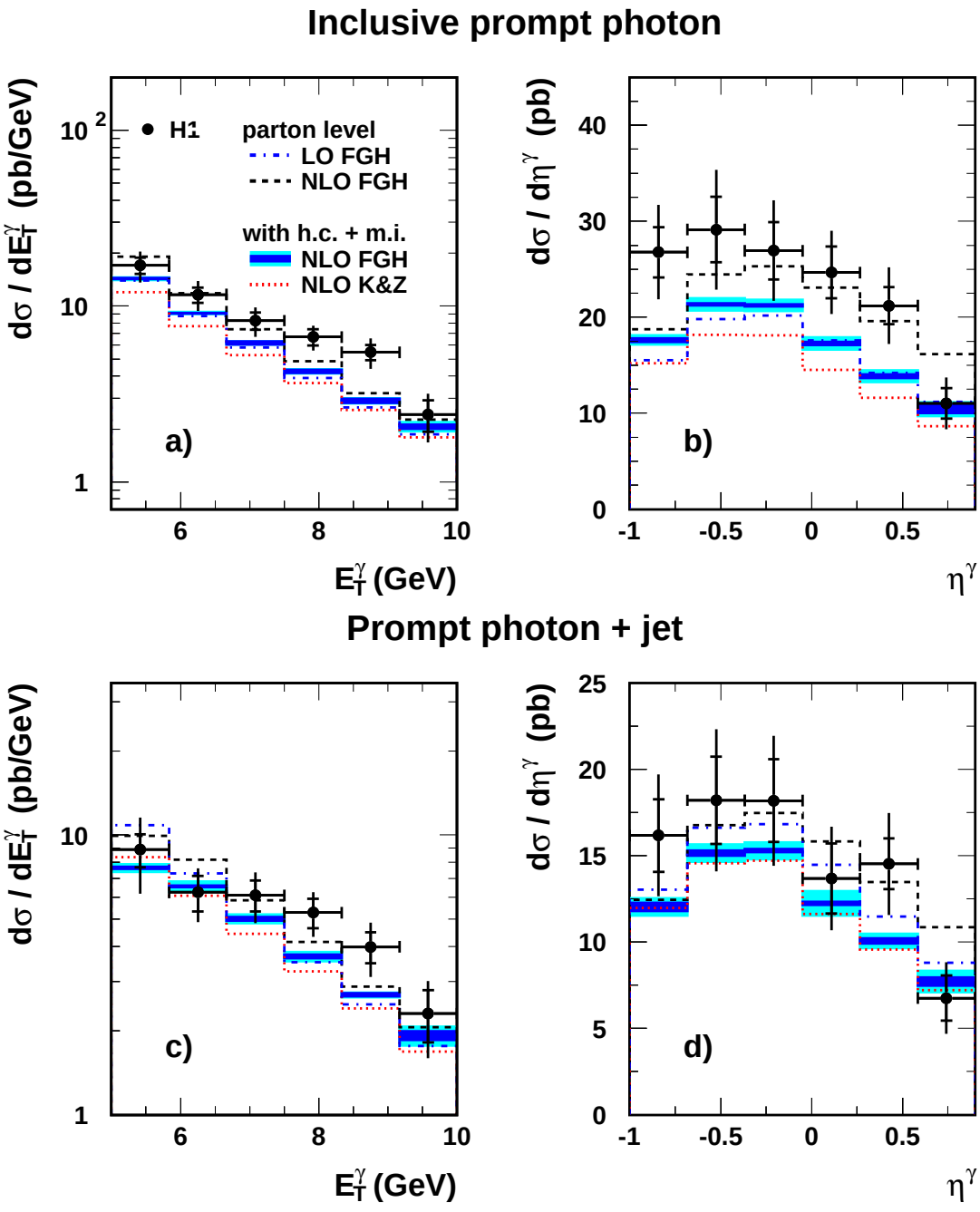

Figure 1.16: Inclusive prompt photon differential cross sections $d\sigma/dE_T^\gamma$ (a) and $d\sigma/d\eta^\gamma$ (b) and photon + jet differential cross sections $d\sigma/dE_T^\gamma$ (c) and $d\sigma/d\eta^\gamma$ (d) in photoproduction as measured by the H1 [35] collabration compared to QCD calculations. Two sets of calculations are used - FGH [39] (blue line) and K&Z [40] (red dotted line). The leading order cross sections are shown with blue dashed-dotted line. The cross sections without hadronisation and multiparton interaction corrections are presented as dashed line.

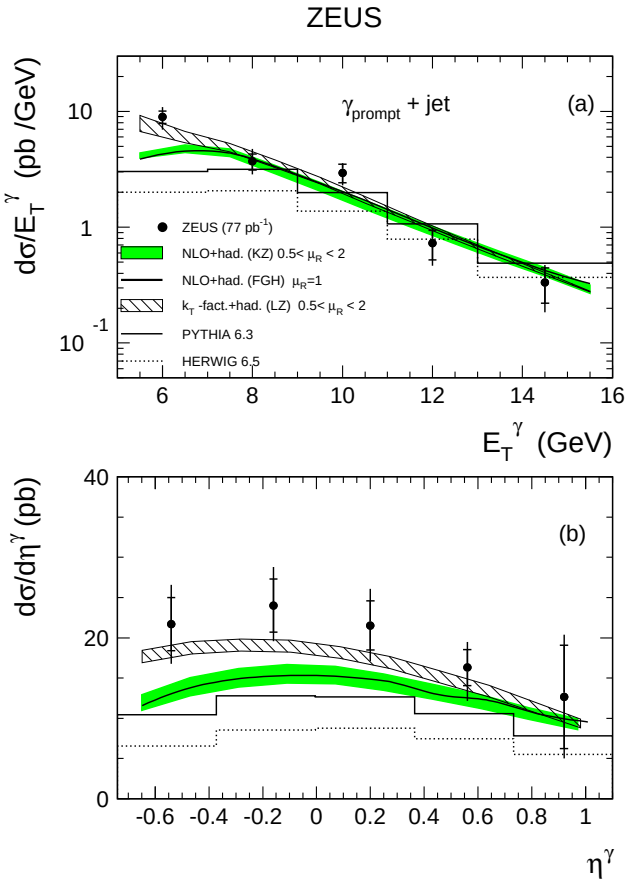

Figure 1.17: Prompt photon + jet differential cross sections $d\sigma/dE_T^\gamma$ (a) and $d\sigma/d\eta^\gamma$ (b) in photoproduction as measured by the ZEUS [38] collaboration compared to predictions by PYTHIA (solid black line), HERWIG (dashed-dotted black line), NLO calculation (green shaded band) and the calculation based on the k_T factorisation approach (striped band).

The importance of the radiation from the electron was shown by the study of the H1 collaboration presented in figure 1.19. The results are quoted for the phase space $4 < Q^2 < 150\,\mathrm{GeV}^2$, $y > 0.05$, $3 < E_T^\gamma < 10\,\mathrm{GeV}$ and $-1.2 < \eta^\gamma < 1.8$ and have been performed using data based on $227\,\mathrm{pb}^{-1}$ integrated luminosity. The H1 measurement in DIS is compared to a LO calculation and to MC prediction, both including radiation from the electron line. The additional contribution of photons radiated from the electron, drastically changes the shape of the η^γ distribution. Nevertheless, both MC and QCD calculation underestimate the prompt photon production rate by roughly a factor two, with the difference being most significant at low Q^2.

The preliminary ZEUS results [42], performed with $320\,\mathrm{pb}^{-1}$ of integrated data luminosity within the kinematic range of $10 < Q^2 < 350\,\mathrm{GeV}^2$, $4 < E_T^\gamma < 15\,\mathrm{GeV}$, $-0.7 < \eta^\gamma < 0.9$ confirm the observations of the H1 collaboration.

1.4.2 Results in hadronic collisions

In hadron collisions prompt photons have been studied for a long time. A large variety of experiments have measured the photon cross section at both fixed target and collider

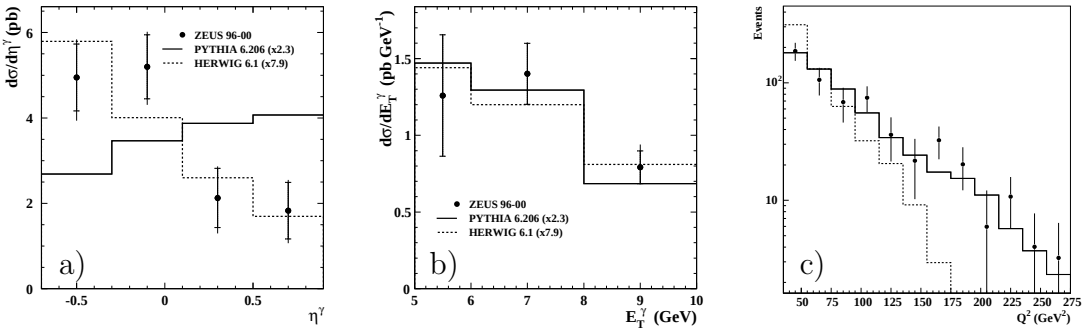

Figure 1.18: Inclusive prompt photon differential cross sections $d\sigma/d\eta^\gamma$ (a), $d\sigma/dE_T^\gamma$ (b) and $d\sigma/dQ^2$ (c) in DIS as measured by the ZEUS [41] collaboration compared to MC predictions.

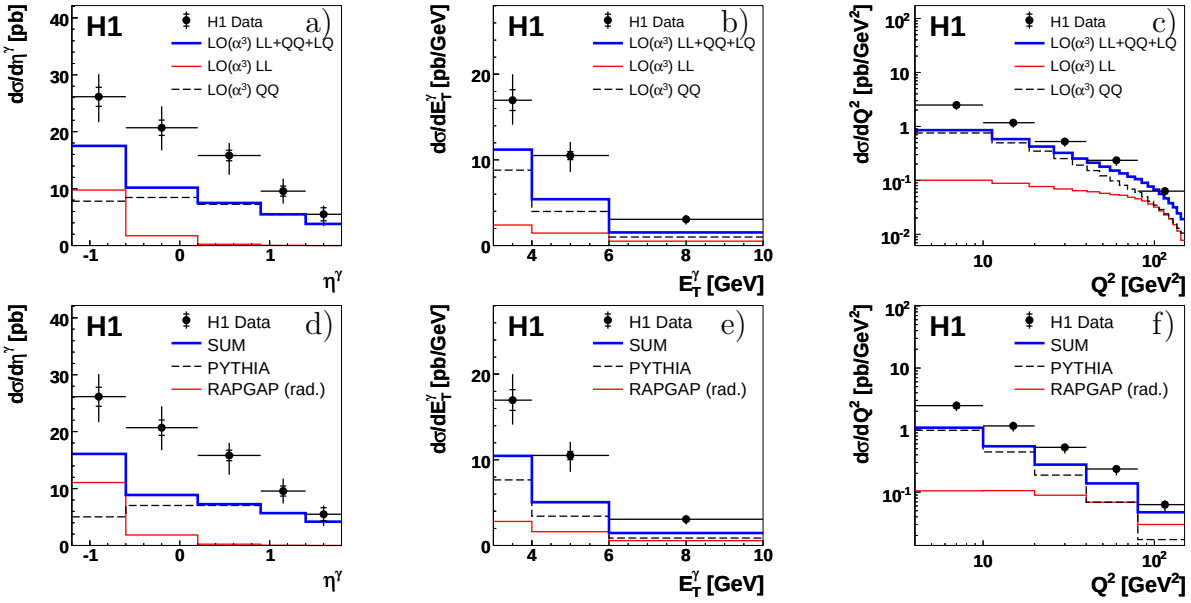

Figure 1.19: Inclusive prompt photon differential cross sections $d\sigma/d\eta^\gamma$ (a, d), $d\sigma/dE_T^\gamma$ (b, e) and $d\sigma/dQ^2$ (c, f) in DIS as measured by the H1 [43] collaboration compared to LO QCD calculation (a-c) and MC prediction (d-f). The direct photon radiation from the quark is indicated by the dashed black line, radiation from the electron line by a red solid line and the sum of the predictions by a blue solid line.

energies. Figure 1.20 summarises results from pp and $p\bar{p}$ collisions. Results are given as the ratio of data to QCD calculations as a function of the reduced variable $x_T = 2E_T^\gamma/\sqrt{s}$. The data used in the figure is summarised in table 1.3 (for further details and references see [44]). With an exception of the measurement performed by E706 experiment, the data agree well with the NLO calculations all over the studied x_T range. The agreement can be still improved by the introduction of 'intrinsic k_T' for the partons within a proton. This situation motivated the introduction of an extra parameter $\langle k_T \rangle$ to account for the effects of multiple soft gluon emissions associated to hard partonic scattering [45]. Authors of [46]

Figure 1.20: Ratio data to theory prompt photon cross section for collider and fixed target hadronic collisions. Figure taken from [44].

1.5 Analysis motivation and goals

Prompt photon results are presented in the form of cross sections with the binning and phase space chosen such as to enhance sensitivity to the studied effects. All the cross sections calculated in this analysis are identified below and shortly discussed. Furthermore, in this section the phase space of the measurement is defined and referred throughout the following chapters.

The production of prompt photons with high transverse momentum can be calculated in perturbation theory, and one may study the prompt photon cross sections in order to test the quality and relevance of the QCD calculations. The measurement of inclusive prompt photon production is performed in bins of its transverse energy E_T^γ and pseudorapidity η^γ. Those measurements are performed in the phase space with the event kinematics conditions and the photon condition later on referred to as the *inclusive* prompt photon phase space. The measurement is restricted to the photoproduction region by the cut on the negative four momentum transfer squared $Q^2 < 1\,\mathrm{GeV}^2$. In addition, to ensure

Experiment	Accelerator	Initial state	$\sqrt{s}$	year
R806	ISR	pp	63 GeV	1982
WA70	SPS	pp	23 GeV	1988
UA1	S$p\bar{p}$S	$p\bar{p}$	630 GeV	1988
R110	ISR	pp	63 GeV	1989
R807	ISR	pp	63 GeV	1990
UA2	S$p\bar{p}$S	$p\bar{p}$	630 GeV	1991
UA6	SppS	$p\bar{p}$	24.3 GeV	1998
UA6	S$p\bar{p}$S	$p\bar{p}$	24.3 GeV	1998
E706	Tevatron fixed target	pBe	31.6 GeV	1998
E706	Tevatron fixed target	pBe	38.8 GeV	1998
D0	Tevatron collider	$p\bar{p}$	630 GeV	2000
D0	Tevatron collider	$p\bar{p}$	1800 GeV	2001
CDF	Tevatron collider	$p\bar{p}$	630 GeV	2001
CDF	Tevatron collider	$p\bar{p}$	1800 GeV	2001

Table 1.3: Summary of the experiments whose results are presented in figure 1.20. The last column indicates the starting year of the given experiment.

that the studied phase space is directly visible by the measurement, a requirement on the inelasticity y is introduced. The low inelasticity region is experimentally dominated by beam gas collision events, while high y events resemble the topology of the deep inelastic scattering. In this analysis, the inelasticity is therefore restricted to $0.1 < y < 0.7$ range. The transverse energy $6 < E_T^\gamma < 15\,\text{GeV}$ and pseudorapidity $-1.0 < \eta^\gamma < 2.4$ of the studied photons is chosen such as to correspond to the acceptance of the experiment. Finally, the measurement is restricted only to photons that are well isolated from the surrounding hadronic activity. As shown in [39, 48] the photon isolation requirement is able to decrease the influence of the quark-to-photon fragmentation, which is not known to a very good precision. Furthermore the isolation requirement is needed for the analysis to suppress background from decay photons. The isolation is ensured with the method similar to the recent prompt photon measurements [38, 43] by the cut on the z variable, $z > 0.9$, where

$$z = \frac{E_T^\gamma}{E_T^{\gamma-jet}}, \tag{1.39}$$

where $E_T^{\gamma-jet}$ is the transverse energy of the jet which the prompt photon is assigned to (for details of the jet algorithm see section 4.2.2). This isolation definition is both collinear and infrared safe.

Collecting the information of both the photon and the leading hadronic jet gives more insight into the underlying dynamics of the prompt photon production process. The most general measurement can be performed in bins of transverse energy and pseudorapidity of both the photon (E_T^γ, η^γ) and the accompanying hadronic jet (E_T^{jet} and η^{jet}). Prompt

Figure 1.21: Photon production cross sections measured in hadronic collisions compared to NLO calculations. The E706 data is scaled by factor 10^{-4}. The figure taken from [44].

photon measurements are also sensitive to the parton density functions (PDFs) of both the proton and the photon and may be used in dedicated analyses to fine tune their parameters. In order to enhance the sensitivity to the parton density functions, the measurement is performed in bins of x_γ^{LO} and x_p^{LO} defined as

$$x_p^{LO} = E_T^\gamma \frac{\exp(\eta^\gamma) + \exp(\eta^{jet})}{2E_p}, \tag{1.40}$$

$$x_\gamma^{LO} = E_T^\gamma \frac{\exp(-\eta^\gamma) + \exp(-\eta^{jet})}{2yE_e}, \tag{1.41}$$

where E_p and E_e are the energies of the incoming proton and electron respectively. Both

variables are estimators of the longitudinal momentum fractions carried by the parton which enter the hard interaction from the proton side x and from the photon side x_γ, already introduced in section 1.2. These leading order definitions of the estimators do not use the energy of the jet. The latter is expected to be reconstructed with relatively low resolution and thus the leading order estimators yield a higher correlation between the generated and reconstructed levels. The usage of the variable x_γ^{LO} is recommended in [49, 50] to reduce infrared sensitivity. The variable x_p^{LO} is discussed e.g. in [51]. The accompanying hadronic jet is required to be within the pseudorapidity range of $-1.3 < \eta^{\text{jet}} < 2.3$ and has the transverse energy above 4.5 GeV. The asymmetric energy cut on the photon and the jet is required to ensure the validity of the NLO calculation [49,52,53] The prompt photon phase space with additional jet requirement is later on referred to as *exclusive* or *photon+jet* phase space.

Another interesting measurement is the study of the transverse correlations between the prompt photon and the accompanying jet. For direct events in leading order and at sufficiently low Q^2, the measured acoplanarity is directly sensitive to the transverse momentum k_T of the parton entering the hard interaction. The mentioned effect can be studied using two variables $\Delta\Phi$ and $p_\perp$ defined as

$$p_\perp \equiv \frac{|\vec{p}_T^{\,\gamma} \times \vec{p}_T^{\,jet}|}{|\vec{p}_T^{\,jet}|} = E_T^\gamma \cdot \sin\Delta\Phi. \tag{1.42}$$

and sketched in the transverse plane[5] in figure 1.22.

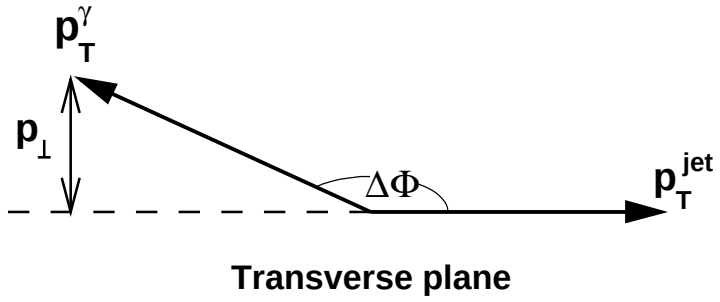

Figure 1.22: Definition of the variables $\Delta\Phi$ and $p_\perp$ used to study the correlations between prompt photon and the leading jet.

The azimuthal acoplanarity between the photon and the jet $\Delta\Phi$ and the photon momentum perpendicular to the jet direction $p_\perp$ were chosen such as to avoid usage of the jet energy measurement and thus minimising the potential degradation of the resolution. If the jet perfectly balances the photon candidate, $p_\perp$ is equal to zero and $\Delta\Phi$ is equal to

[5]For a definition of the transverse plane the reference system is needed. It is introduced in section 3.2 where transverse plane is also defined.

Inclusive Phase Space	**Exclusive Phase Space**
$Q^2 < 1\,\mathrm{GeV}^2$	$Q^2 < 1\,\mathrm{GeV}^2$
$0.1 < y < 0.7$	$0.1 < y < 0.7$
$z > 0.9$	$z > 0.9$
$6.0 < E_T^\gamma < 15\,\mathrm{GeV}$	$6.0 < E_T^\gamma < 15\,\mathrm{GeV}$
$-1 < \eta^\gamma < 2.4$	$-1 < \eta^\gamma < 2.4$
	$E_T^{\mathrm{jet}} > 4.5\,\mathrm{GeV}$
	$-1.3 < \eta^{\mathrm{jet}} < 2.3$

Table 1.4: The inclusive and exclusive prompt photon phase space definitions.

Direct Enhanced Phase Space	**Resolved Enhanced Phase Space**
$Q^2 < 1\,\mathrm{GeV}^2$	$Q^2 < 1\,\mathrm{GeV}^2$
$0.1 < y < 0.7$	$0.1 < y < 0.7$
$z > 0.9$	$z > 0.9$
$6.0 < E_T^\gamma < 15\,\mathrm{GeV}$	$6.0 < E_T^\gamma < 15\,\mathrm{GeV}$
$-1 < \eta^\gamma < 2.4$	$-1 < \eta^\gamma < 2.4$
$E_T^{\mathrm{jet}} > 4.5\,\mathrm{GeV}$	$E_T^{\mathrm{jet}} > 4.5\,\mathrm{GeV}$
$-1.3 < \eta^{\mathrm{jet}} < 2.3$	$-1.3 < \eta^{\mathrm{jet}} < 2.3$
$x_\gamma^{LO} > 0.8$	$x_\gamma^{LO} < 0.8$

Table 1.5: The direct and resolved enhanced prompt photon phase space definitions.

180°. Both variables are studied in the phase space with enhanced contribution from either direct or resolved events with x_γ^{LO} variable used to split the exclusive phase space accordingly. The part of the phase space with enhanced contribution of direct events ($x_\gamma^{LO} > 0.8$) is referred later on as the *direct* phase space. The complementary part ($x_\gamma^{LO} < 0.8$), with enhanced contribution from the resolved events is referred to as the *resolved* phase space.

All the phase spaces defined in this section are summarised in tables 1.4 and 1.5.

Chapter 2

Theoretical predictions

The following chapter describes theoretical predictions to the prompt photon production studied in this analysis. The Monte Carlo models are described in section 2.1. The pQCD calculations performed for a direct comparison with the measurement are introduced in section 2.2.

2.1 Monte Carlo simulation

The Monte Carlo [1] (MC) term describes a broad set of problem solving methods based on a stochastic approach with pseudorandom numbers statistically probing the problem. They are particularly useful for problems that cannot be solved analytically, like calculation of multi-dimensional integrals with complicated boundary conditions. MC techniques are also fit for studying complex systems with a large number of possibly coupled degrees of freedom. The flexibility of MC makes it useful in a broad area which ranges from computational physics, chemistry, biochemistry, climate sciences through to business applications and risk management simulations.

In high energy physics MC simulations are extensively used for modelling hard interactions. They have the advantage of including non-perturbative processes such as fragmentation which cannot be calculated in QCD and need to be modelled. MC simulations are also used to study various detector effects and to determine corresponding detector corrections.

2.1.1 MC event generators

MC event generation is divided into a set of successive steps, as illustrated in figure 2.1 for the case of ep scattering:

[1] The history of the name can be traced back to the Manhattan Project in Los Alamos Scientific Laboratory where two physicists, John von Neumann and Stanisław Ulam proposed for the first time a stochastic method to be applied to solve the problem that could not be solved by theoretical calculations. Being secret, their project got the code name after the famous casino in Monaco.

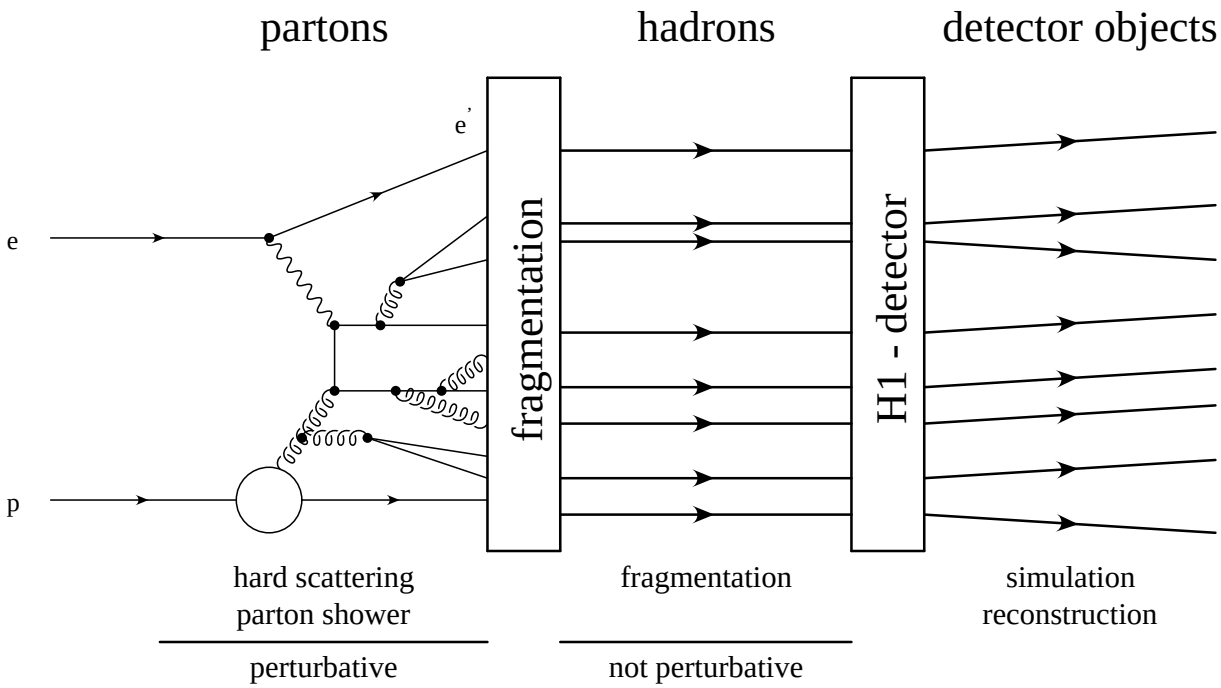

Figure 2.1: MC generation process in the *ep* scattering case. The figure is taken from [54].

- **Hard scattering**
  The interaction between particles are described by matrix elements calculated according to the Feymann rules. This perturbative step determines the main characteristics of the event. The parton from the proton is chosen according to parton density functions (PDFs) describing the probability of finding a parton in a proton at a given Q^2 and x values. The photon PDF is used for choosing a parton from the photon in case of resolved events. For more details concerning PDFs see section 1.2.

- **Initial and final state radiation**
  The existence of charged or coloured objects before or after hard interaction may lead to large corrections due to photon or gluon emissions. The electromagnetic radiation is modelled according to QED.

  QCD corrections may be modelled by the so called Parton Shower (PS) model [55] describing the parton cascade by splitting of a parent parton into two daughters. The possible transitions are $q \to qg$, $g \to q\bar{q}$, $g \to gg$ and the evolution of the whole shower is based on the DGLAP equations (see section 1.2) and splitting functions $P_{\alpha\beta}(x/z)$. The PS model can be applied for both initial and final state radiation and therefore the parton entering the hard interaction may already originate from a parton splitting.

  An alternative to the PS model is given by a colour dipole model (CDM) [56] where each pair of coloured objects is treated as a colour dipole emitting gluons. The radiation leads to additional dipoles emitting gluons and resulting in a parton cascade. In case of *ep* scattering, the cascade is initiated by the dipole constructed

by a struck quark and the proton remnant.

- **Fragmentation**
  After the radiation step, as distances between partons increase, QCD becomes strongly interacting and the perturbation theory breaks down. The coloured partons are combined into colourless hadrons in a process called fragmentation. Since the fragmentation region is non perturbative, some phenomenological models are used.

  In the Lund string model [57, 58], two colour charged objects are bound together forming a string. As two particles move apart, the potential energy of the string increases until enough energy is accumulated to produce a quark-antiquark pair in a break up of the string. Newly created quarks get connected with the loose ends of the original string and the process continues until stable hadrons emerge.

  An alternative to the Lund string model is provided by the cluster decay fragmentation (CDF) model [59]. In this approach all gluons produced in parton showers are decayed into quark-antiquark pairs. Colourless pairs of $q\bar{q}$ form clusters consequently decaying into final state hadrons.

- **Detector simulation**
  In the last step, the final state particles are passed through the detector simulation producing a comparable output with information collected in real data experiment. In the case of the H1 experiment (see chapter 3), the H1SIM [60] software package, based on the GEANT [61] program provides a detailed description of the H1 detector including all the instrumentation as well as passive material. Due to the high time consumption of the detailed shower simulation in the LAr calorimeter [62, 63] (see section 3.2.2), it has been exchanged by the significantly faster H1FAST [64, 65] shower parameterisation. It has been proven [66] that the used parameterisation gives reliable shower description even in such a detailed shower shape analysis (see chapter 6) as used for this thesis. In the same way as for real data, the output of the H1SIM package is passed through the H1REC [67] reconstruction program, which interprets the electronic signals and creates composed objects.

As pictured in figure 2.1, the event properties can be studied using partons, hadrons or detector objects and are referenced later on as parton, hadron and detector levels. The parton (generator) level takes into account all the particles after the radiation step. The hadron level uses all the stable final state particles produced in the fragmentation step as well as stable particles that do not hadronize. Finally, experimentally measured detector level uses all the objects obtained after the reconstruction stage, such as tracks and clusters. The direct comparison between data and theory is possible only on the same level, usually being the hadron level. MC, having its representation on all three levels is used for transition between different levels.

2.1.2 MC samples

More than thirty different MC samples were generated for their specific purposes in this analysis. In this subsection the ones directly involved in the calculation of the results are shortly described. Four different generators were used for the MC production:

- **PYTHIA**
  The main MC event generator used to account for the majority of the effects is the PYTHIA [68] program (version of 6.224). The event generation is based on a LO matrix elements for quarks scattering with a transverse or longitudinal polarised virtual photon: $q\gamma_T^* \rightarrow q\gamma$ and $q\gamma_L^* \rightarrow q\gamma$. PYTHIA computes parton showers with the PS method and uses the JETSET [69] program to model Lund string fragmentation. Multi parton interactions (see section 8.7.1) are implemented by adding interactions between spectator partons within the same event.

- **HERWIG**
  The HERWIG [70] program (used in version 6.505), in contrast to PYTHIA is using the CDF model to describe fragmentation and thus it is used to evaluate model dependence of the hadronisation corrections (see section 8.7.3).

- **RAPGAP**
  The RAPGAP MC event generator [71] was used to simulate the background from low Q^2 DIS events and to study its influence on the selected data sample. Similarly to PYTHIA it uses initial and final state PS and the Lund string fragmentation model.

- **DJANGOH**
  The DJANGOH MC event generator [72] was used in order to generate high Q^2 DIS background. The sample generated by DJANGOH complements DIS phase space generated by RAPGAP.

The bulk of the MC events was generated by PYTHIA running in the prompt photon channel with direct and resolved subprocess codes (ISUB) listed in table 2.1. In addition, the radiated photons, both in the direct and resolved mode are selected from a di-jet MC sample, produced with subprocesses given in table 2.2 with a hard radiated photon in the final state. The high statistics of 11 milion events is needed to sufficiently populate the unfolding matrix (see section 8.2.1). The sample is referenced as MC set I in table 2.3, which summarises all the MC signal samples.

To sufficiently populate high transverse energies, a prompt photon sample (MC set II) with prescaled low E_T^γ prompt photon events ($E_T^\gamma < 10\,\text{GeV}$) is generated. In order to correctly take into account the changing detector condition, both samples (MC set I and II) were simulated four times for four different running periods additionally increasing statistics of the signal MC. Samples were generated with CTEQ6L [73] serving as a proton PDF and SASG-1D [74] as a photon PDF.

Direct		Resolved	
ISUB	Process	ISUB	Process
34	$f_i\gamma \to f_i\gamma$	14	$f_i\bar{f}_i \to g\gamma$
		18	$f_i\bar{f}_i \to \gamma\gamma$
		29	$f_ig \to f_i\gamma$
		114	$gg \to \gamma\gamma$
		115	$gg \to g\gamma$

Table 2.1: Processes included in the PYTHIA signal MC generation.

Direct		Resolved	
ISUB	Process	ISUB	Process
33	$f_i\gamma \to f_ig$	11	$q_iq_j \to q_iq_j$
54	$g\gamma \to f_k\bar{f}_k$	12	$q_iq_i \to q_kq_k$
		13	$q_iq_i \to gg$
		28	$q_ig \to q_ig$
		53	$gg \to q_kq_k$
		68	$gg \to gg$

Table 2.2: Processes included in the PYTHIA radiated signal and background MC generation.

Parton level pQCD calculations (see section 2.2) need to be corrected to the hadron level (corrections themselves are explained in section 8.7) for a comparison to the measurement. For this specific purpose and to study multiparton interaction (MPI) effect (see section 8.7.1), MC sets III and IV were generated. Since the detector level is not used in this case, the simulation and reconstruction steps were skipped.

Further MC sets were used to evaluate uncertainties of the corrections applied to the pQCD calculations. The uncertainties arise from: multiparton interaction model (MC sets V-VIII), hadronisation model (MC set IX-X) and the choice of PDF sets (MC sets XI-XV). All those effects are explained in section 8.7. Since the photon PDF participates only in the generation of resolved events, direct event samples are not needed in every PDF set configuration. As for MC set III, the simulation and reconstruction steps were skipped.

Table 2.4 summarises all the MC samples used to study the background to the analysis presented in this thesis. The main background arises from neutral hadrons decaying into multiple photon final state (see chapter 6). To properly take it into account, a huge photoproduction MC sample (MC set XVI) with typical photoproduction processes (summarised in table 2.2) was generated. Due to analysis cuts optimised for background

rejection, its selection rate is of the order 10^{-5}. In order to avoid processing of not needed events, only events with an isolated neutral particle with transverse momentum $p_T > 4\,\text{GeV}$ are accepted for simulation already on the generator level. To avoid double counting, events with a hard radiated photon are removed from the sample. Both requirements remove 98% of all generated events. In the remaining sample, the final selection rate for direct events amounts to $\sim 1.0\%$, while for resolved events it drops to $\sim 0.5\%$. From the originally generated 0.8×10^9 direct background events and 1.5×10^9 resolved background events only approximately 170×10^3 direct and 140×10^3 resolved events were used for the discrimination process described in chapter 6.

The DIS background was studied using three different MC samples: prompt photon sample produced in low Q^2 DIS events (MC set XVII), photons radiated from an electron line in low Q^2 DIS events (MC set XVIII) and high Q^2 events (MC set XIX). The low Q^2 DIS prompt photon events were studied in [43] and the same signal and background samples were used in this analysis. DIS events may migrate into the selection if scattered electron in not correctly identified, or in case of high Q^2 events, it is misidentified as photon.

The shower shape analysis, described in chapter 6, requires additional statistically independent signal and background samples. Since only the LAr calorimeter signal initiated by a single photon and photons coming from decayed isolated neutral hadrons is needed, single particle (SP) MC was produced. SP photons (MC set XX) and SP hadrons (MC set XXI-XXX) were generated flat both in transverse energy and pseudorapidity and simulated for 2005 running period. Hadron samples were mixed according to the study presented in [66]. Table 2.5 summarises all the SP MC samples.

MC set	Generator	Comments	Simulation period	Luminosity $[pb^{-1}]$				N unweighted events $[\times 10^3]$			
				dir.	res.	drad.	rrad.	dir.	res.	drad.	rrad.
MC Signal				(simulated $\times 4$)				(simulated $\times 4$)			
I	PYTHIA		04/05/06/07	71958	110233	16940	3443	5405	5696	110	281
MC Signal (high E_T^γ)				(simulated $\times 4$)				(simulated $\times 4$)			
II	PYTHIA		04/05/06/07	7442	10333	–	–	254	221	–	–
Correction to the calculations											
III	PYTHIA	∼MPI	–	6202	8615	4978	1256	584	406	31	90
IV	PYTHIA	MPI	–		5162	–	2092	–	262	–	150
Multiparton interaction correction uncertainty											
V	PYTHIA	$p_T^{min} = 2.2\,\text{GeV}$ ∼MPI	–	6199	8619	–	–	585	399	–	–
VI	PYTHIA	$p_T^{min} = 2.2\,\text{GeV}$ MPI	–	–	8621	–	–	–	404	–	–
VII	PYTHIA	$p_T^{min} = 1.6\,\text{GeV}$ ∼MPI	–	6200	8594	–	–	584	397	–	–
VIII	PYTHIA	$p_T^{min} = 1.6\,\text{GeV}$ MPI	–	–	8596	–	–	–	376	–	–
Hadronisation correction uncertainty											
IX	HERWIG		–	14911	20334	–	–	2132	1664	–	–
X	HERWIG	∼MPI	–	7794	11038	–	–	725	520	–	–
PDF uncertainty											
XI	PYTHIA	$\text{p}_{PDF}^{CTEQ5L} \times \gamma_{PDF}^{AFG}$	–	1222	8597	–	–	128	433	–	–
XII	PYTHIA	$\text{p}_{PDF}^{CTEQ5L} \times \gamma_{PDF}^{GRV}$	–	–	8593	–	–	–	435	–	–
XIII	PYTHIA	$\text{p}_{PDF}^{CTEQ6L} \times \gamma_{PDF}^{AFG}$	–	1240	5154	–	–	126	260	–	–
XIV	PYTHIA	$\text{p}_{PDF}^{MRST04} \times \gamma_{PDF}^{AFG}$	–	1136	8595	–	–	126	433	–	–
XV	PYTHIA	$\text{p}_{PDF}^{MRST04} \times \gamma_{PDF}^{GRV}$	–	–	8594	–	–	–	433	–	–

Table 2.3: Summary of direct (dir.), resolved (res.), radiated direct (drad.) and radiated resolved (rrad.) signal MC samples.

MC set	Generator	Comments	Simulation period	Luminosity $[pb^{-1}]$ dir.	Luminosity $[pb^{-1}]$ res.	N unweighted events $[\times 10^3]$ dir.	N unweighted events $[\times 10^3]$ res.
Photoproduction background							
XVI	PYTHIA		05	45248	5244	15584	29902
Low Q^2 DIS background							
XVII	PYTHIA	$Q^2 > 3\,\mathrm{GeV}^2$, prompt-γ	05	1056	–	193	–
XVIII	RAPGAP	$Q^2 > 3\,\mathrm{GeV}^2$, inclusive	05	49	–	540	–
High Q^2 DIS background							
XIX	DJANGOH	$Q^2 > 100\,\mathrm{GeV}^2$, inclusive	05	655	–	3960	–

Table 2.4: Summary of direct (dir.) and resolved (res.) background MC samples.

MC set	Particle	Simulation period	N events $[\times 10^3]$
XX	γ	05	900
XXI	π^0	05	825
XXII	η	05	808
XXIII	η'	05	300
XXIV	ω	05	300
XXV	n	05	300
XXVI	$\bar{n}$	05	300
XXVII	K_L^0	05	300
XXVIII	K_S^0	05	300
XXIX	ρ	05	300
XXX	K^*	05	300

Table 2.5: Summary of single particle MC samples.

2.2 QCD calculations

In this section, two QCD calculations used in this analysis are presented: a full NLO collinear approach and a calculation based on the k_T factorisation approach. Calculations are performed in the same phase space as the measurement (listed in tables 1.4 and 1.5), with the sole exception of the isolation criteria (discussed in section 8.7.2).

2.2.1 NLO collinear calculation

The full NLO collinear calculation by Fontannaz, Guillet and Heinrich (FGH) [39, 75] includes the leading order direct and resolved processes $\gamma q \to \gamma q$ and NLO corrections to these terms. Besides the direct hard interaction, photons may also be produced from the fragmentation of a high p_T quark or gluon in the final state (see section 1.3). The fragmentation process, described by a fragmentation function, is included in the calculation, as well as the direct box diagram shown in figure 1.13a. The box diagram is calculated to order α_s^2. The contribution from quark to photon fragmentation to the total cross section is at the level of 4%, while box diagram amounts to roughly 10% of the total cross section. The calculation uses the parton density functions CTEQ6L [73] for the proton and AFG04 [50] for the photon. The scales for renormalisation μ_R and factorisation μ_F, are chosen to be $\mu_R = \mu_F = E_T^\gamma$. The NLO corrections to the LO cross section are significant in the inclusive sample. They increase the predicted cross section by factor between $1.1 - 1.4$, being largest at low E_T^γ and large η^γ (the detailed study of NLO correction is presented in section 9.3). For the exclusive phase space, the corrections are much smaller and below 10% on average.

2.2.2 $k_\perp$ factorisation calculation

Conventional parton densities are defined in terms of an integral over all transverse momenta and virtuality for a parton that initiates hard scattering. While such a definition of an integrated parton density is appropriate for very inclusive quantities, such as ordinary structure functions F_1 and F_2 in DIS, the definition becomes increasingly unsuitable as one studies less inclusive cross sections, where introduced approximations on parton kinematics become visible. The next step of PDF development is the usage of k_T unitegrated parton densities functions (uPDF), where k_T denotes the original transverse momentum of the parton from the composed particle. The calculation that convolutes uPDF with the matrix elements are called k_T-factorisation [76, 77] based calculations (for more complete review of the method see [78, 79]).

The calculation of Lipatov and Zotov (LZ) [80] are based on the $k_\perp$ factorisation approach. The calculation uses the unintegrated quark and gluon densities of the photon and the proton using the Kimber-Martin-Ryskin (KMR) prescription [81] with the GRV parameterisations of the colinear quark and gluon densities [82, 83]. The k_T-factorisation approach is expected to account for the main part of the colinear higher order QCD corrections [80]. Direct and resolved processes are considered in the calculation while the

contributions from fragmentation and from the box diagram are neglected.

Chapter 3

The H1 experiment at HERA

This chapter introduces the experimental apparatus used to collect the data that this analysis is based on. At first, in section 3.1, an overview of the HERA[1] *ep* collider is given. The most important elements of the H1 detector are introduced in section 3.2. Special care is taken to describe parts particularly relevant for this analysis.

3.1 HERA storage ring

The HERA [84] storage ring in located at the DESY[2] research facility in Hamburg, Germany. It is an unique electron[3]-proton collider providing an opportunity to study lepton-parton interaction at high energies. Figure 3.1 shows an overview of the HERA facility.

The proton beam preparation starts with the acceleration of negatively charged hydrogen ions up to 50 MeV in the linear accelerator (H^--linac). In order to obtain the proton beam, ions are stripped off the electrons before the injection into the DESY III ring. Seventy *bunches* with a bunch crossing distance of 96 ns, identical to HERA, are accelerated up to 7.5 GeV in the DESY III ring. The bunches are then transferred into the PETRA[4] storage ring, where they are accelerated further. The 70 GeV protons are injected to the HERA ring. The procedure is repeated until HERA is filled with 210 proton bunches. In the HERA ring, protons are accelerated to their final energies of 920 GeV.

The lepton acceleration starts with the e^--linac (e^+-linac in case of positron beam) accelerator, where they are accelerated to the energies of 450 MeV. A single bunch of electrons ≈ 60 mA is transferred into the DESY II ring, where it is accelerated to the energies of 7 GeV. Seventy electron bunches are injected into the PETRA II ring, where they reach the energy of 14 GeV. The final energy of 27.6 GeV is achieved at the last stage of the acceleration in the HERA ring, where 210 electron bunches circle the ring with 96 ns

[1] **H**adron **E**lectron **R**ing **A**nlage.

[2] **D**eutsches **E**lektronen **Sy**nchrotron.

[3] Unless otherwise stated, here and in the following the term electron refers to both the electron and the positron.

[4] **P**ositron **E**lektron **T**andem **R**ing **A**nlage

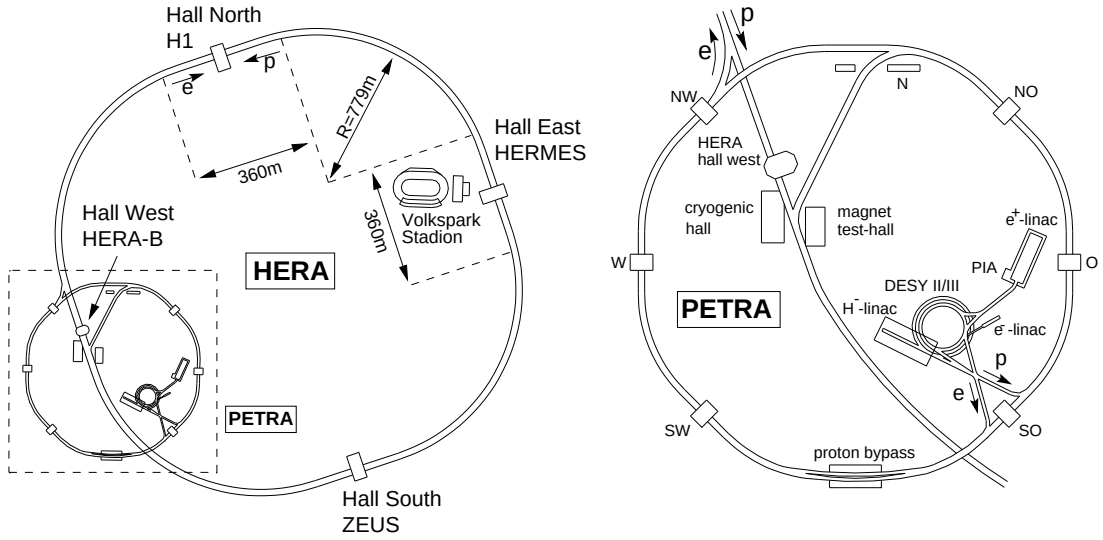

Figure 3.1: Layout of the HERA collider (left) and its pre-accelerator system (right).

spacing.

With the rate of 10.4 MHz, defining the *HERA-clock* electron and proton bunches cross each other. The head-on collisions of 27.6 GeV electrons and 920 GeV protons lead to the system with centre-of-mass energy of $\sqrt{s} = 319\,\text{GeV}$ (see equation 1.3).

There are four experiments located on the HERA ring. Two general purpose experiments, H1 [85] and ZEUS [86] are detectors working on *ep* head-on collisions. The HERMES experiment [87] records the scattering of polarised electrons on polarised gas targets and has been commissioned to study the spin structure of nucleons. The HERA-B experiment [88] was designed to study the CP violation in B^0-B^0 system generated through collisions of beam protons with a stationary target.

The performance of the accelerator is characterised by the produced integrated luminosity L which is proportional to the number of events expected in the scattering experiment N^{ep},

$$N^{ep} = L \cdot \sigma^{ep}, \tag{3.1}$$

where σ^{ep} is the total *ep* scattering cross section. The integrated luminosity is defined as the time-integrated instantaneous luminosity $\mathfrak{L}$,

$$L = \int \mathfrak{L} dt. \tag{3.2}$$

$\mathfrak{L}$ depends on the bunch crossing frequency, bunch currents and the collimation of the beams. Figure 3.2 presents the integrated luminosity collected by the H1 experiment during two HERA running periods. The HERA I running period extends between the first collisions recorded in 1992 and the year 2000, when HERA went through the major improvement effort [89]. After the upgrade the instantaneous luminosity is increased compared to the previous conditions by approximately a factor five to $\mathfrak{L}^{HERAII}_{designed} = 7.4 \cdot 10^{31}\ \text{cm}^{-2}\text{s}^{-1}$. The HERA II running period counts between years 2004 and 2007 when

the HERA collider has been switched off. Last few months of the HERA II running were dedicated for the special collision with decreased proton energy to 460 GeV and 575 GeV. The decrease of the proton current is reflected in the lower slope of the integrated luminosity plot at the end of the HERA II run.

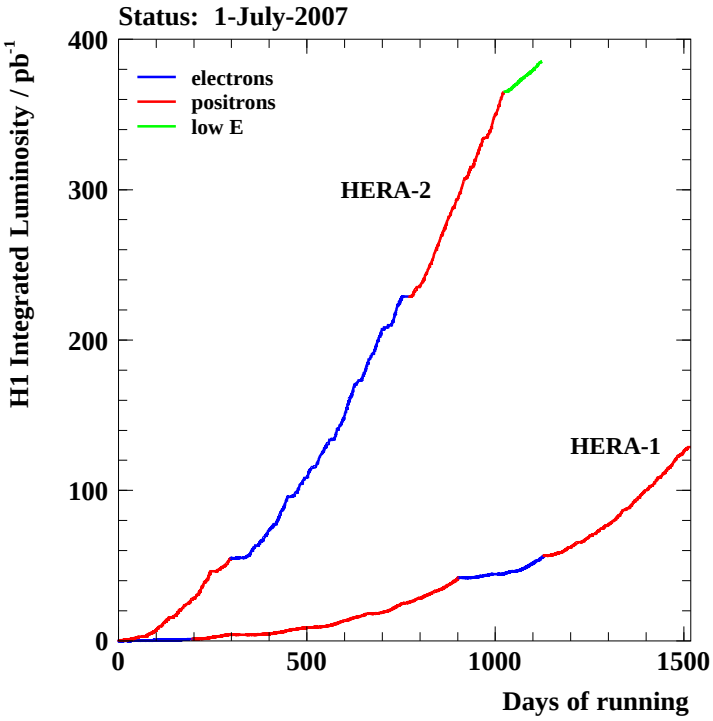

Figure 3.2: HERA I and HERA II integrated luminosity as a function of time.

3.2 H1 detector

The multi-purpose detector H1 was designed to measure the final state particles resulting from the *ep* collisions provided by the HERA machine. Its complex set of subdetectors build around the nominal interaction point IP allows precise identification and reconstruction of particles emerging from the interaction. It provides almost hermetic geometrical coverage, with the main limitation being the beam pipes.

Figure 3.3 presents the H1 detector drawn with its major components highlighted. The IP is surrounded by the tracking system consisting of a variety of subdetectors. The innermost system, the tracking systems is enclosed by the calorimetry dedicated to measure energy of the final state particles. Both tracking and calorimetry systems are contained within the homogeneous magnetic field of 1.15 T provided by superconducting magnets. The beam energy asymmetry is reflected in the geometry of the detector. The proton beam direction, referred to as the *forward* direction is significantly heavier instrumented. The opposite, electron beam, or *backward* direction is dedicated predominantly to the reconstruction of the scattered electron.

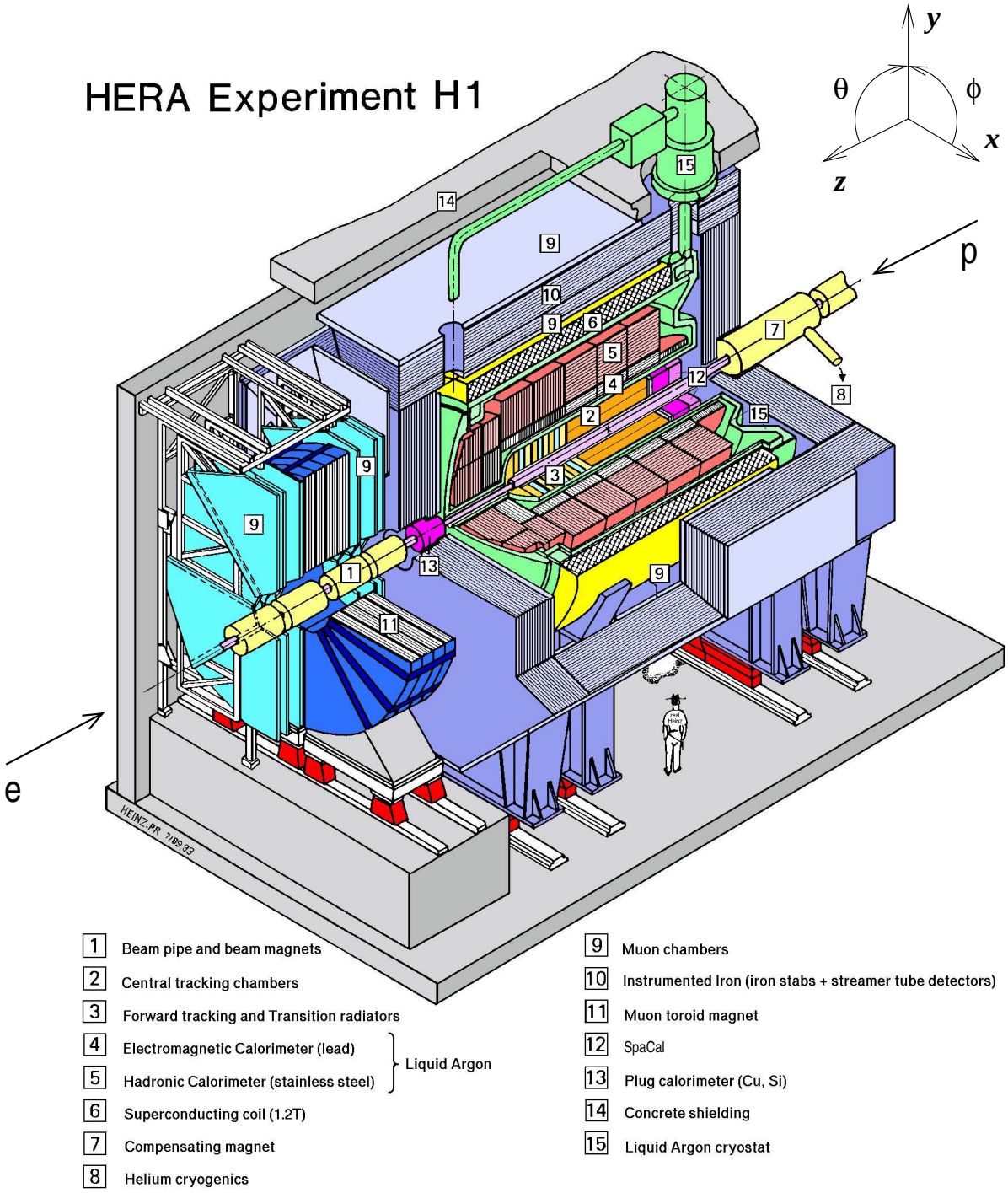

Figure 3.3: The H1 detector with its major components and the reference system.

The right-handed reference frame at H1 is defined such that the positive z-axis points into the proton direction, positive x-axis point horizontally to the centre of the storage rings and positive y-axis points vertically upwards. The polar angle θ is measured with respect to the z-axis and the azimuthal angle ϕ lies in the xy-plane, referred to as a *transverse* plane, with $\phi = 0$ indicating the x-axis. The radial distance r lays also in the

transverse plane and is defined as $r = \sqrt{x^2 + y^2}$. The reference system is presented in the upper-right corner of the figure 3.3. The *pseudorapidity* η, defined as

$$\eta = -\ln \tan(\theta/2) \tag{3.3}$$

is valued for its invariance[5] under longitudinal Lorentz-boosts and is often used instead of the polar angle θ.

The detailed description of the H1 detector can be found in [67, 85]. Here, only the most relevant for the present work subdetectors are briefly described.

3.2.1 Tracking

The interaction point is tightly enclosed by the H1 tracking system, constituting the innermost part of the detector. Figure 3.4 gives a longitudinal view of the system with its most important components indicated. The H1 tracking system consists of the set of drift chambers, multi-wire proportional chambers and silicon trackers and is divided into the central tracking detector (CTD) and the forward tracking detector (FTD).

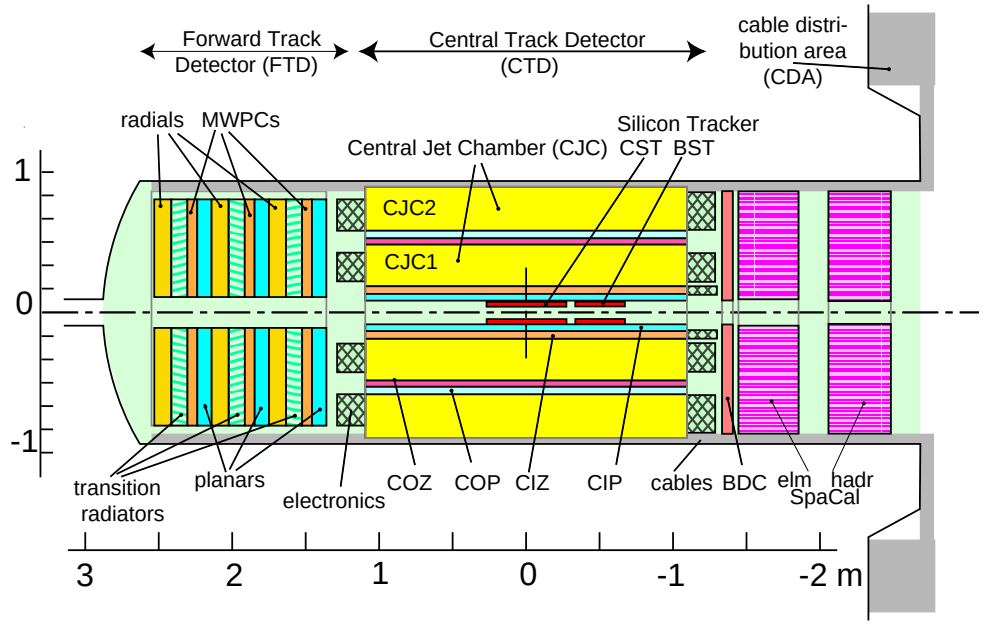

Figure 3.4: A longitudinal view of the H1 tracking system including the central and forward tracking detectors.

The main part of the CTD has an active length of 220 cm along the z direction and consist of two concentric drift chambers, the inner CJC1 ($20.3 < r < 45.1$ cm) and the outer CJC2 ($53.0 < r < 84.4$ cm) [90]. The CJC1 (CJC2) is azimuthally segmented into 30 (60) drift cells with 24 (32) sense wires per cell aligned parallely to the z direction. Charged particles travelling through the chambers are traced by their ionising of gas atoms[6]. The spatial

[5] Strictly speaking, only the difference in pseudorapidity is invariant.
[6] In the CJC the mixture of Ar, CO_2 and CH_4 is used.

resolution in the $r\phi$ plane is based on a drift-time measurement of the ionisation electrons, while the z-coordinate is obtained from charge division. The *hit* position is measured with an accuracy of 170 μm in the $r\phi$ plane and 3 cm along the z-axis. The z resolution can be improved by employing the *Central Inner Proportional* (CIP) chamber [91] and *Central Outer Z* (COZ) chamber [92], which leads to an improvement of the z measurement resolution up to typically 300μm.

The CIP chamber (see figure 3.5) is a multiwire proportional chamber (MWPC) which consist of five cylindrical detector layers with cathode pad readout. It is located between $15 < r < 20$ cm and has a total active length of 2.2 m. The chamber is segmented into 16 azimuthal sectors and 93-119 (depending on the layer) pads along the z direction. Multiwire proportional chambers are characterised by a very fast response to ionising particles. A typical value for intrinsic time resolution is 10 ns. and for that reason, the CIP chamber provides the information used already on the level one trigger level (see section 3.2.5).

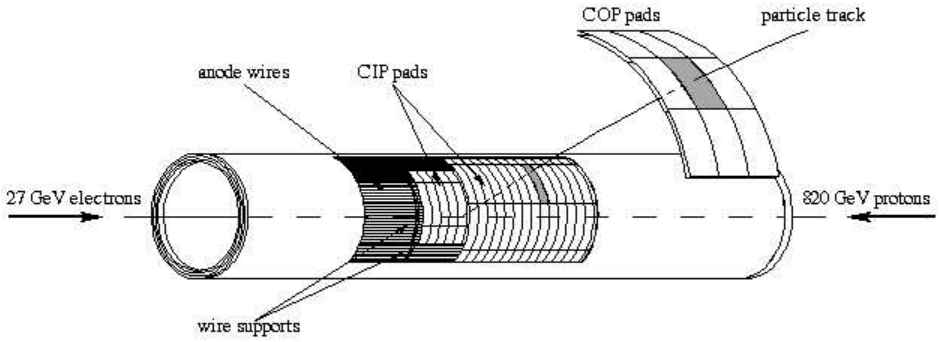

Figure 3.5: The structure of the CIP chamber.

The FTD [93] is a set of drift chambers designed to detect forward tracks in the polar angle range $5° < \theta < 25°$. It consists of three identical supermodules aligned along the z-axis. Each of the supermodules contains a planar drift chamber and a proportional chamber.

The Backward Proportional Chamber (BPC) [94] covers the polar angle in range $153° < \theta < 174°$ and is used to enhance the angle measurement of the scattered electron at low Q^2.

The innermost part of the tracking system is composed of the silicon detectors, the Backward Silicon Detector (BST), the Central Silicon Detector (CST) and the Forward Silicon Detector (FST). The BST [95] has an acceptance coverage of $164° < \theta < 176°$ and a resolution $\sigma_r = 22$ μm in r-direction. It is used to identify the scattered electron at small polar angles. The CST [96] consists of two layers of silicon strip detectors, arranged parallel to the beam pipe. The polar angle acceptance range is $30° < \theta < 150°$ and the resolution is $\sigma_z = 12$ μm in z and $\sigma_{r\phi} = 22$ μm in both r and ϕ. It is used for a precise reconstruction of the interaction vertex. The FST [97] consists of five planes of silicon strip detector covering the polar angle $8° < \theta < 16°$ and having resolution similar to the BST.

3.2.2 Calorimetry

Calorimeters are detector components providing the energy measurement for both charged and neutral particles. H1 calorimetry system consists of four separate calorimeters, with the main one, being the Liquid Argon (LAr) calorimeter covering the forward and central region of the detector. Energy leaking out of the LAr is measured by the tail catcher installed in the instrumented iron. The plug calorimeter and Spaghetti calorimeter (SPACAL) provide energy measurements in the forward and backward detector regions respectively.

The LAr calorimeter [98] is the most important subdetector used in the current analysis. It provides the identification of prompt photons as well as measures energy of hadronic final state particles. Due to its vital importance it is described below in some more details.

Figure 3.6 shows a longitudinal cross section of the LAr calorimeter. It is geometrically divided into eight wheels, listed from the backward direction these are the Backward Barrel (BBE), three Central Barrels (CB1, CB2 and CB3), two Forward Barrels (FB1 and FB2) and Inner and Outer Forward module (IF and OF). Except of BBE, which is purely electromagnetic and OF, which is purely hadronic, all the wheels are equipped with electromagnetic and hadronic sections. In addition to the wheel segmentation, LAr calorimeter is divided into eight octants along the azimuthal angle ϕ. Between octants and wheels, there are insensitive calorimeter regions, which are referred later on as to ϕ-cracks and z-cracks, respectively. The LAr has an asymmetric polar angle coverage of $4^\circ < \theta < 154^\circ$.

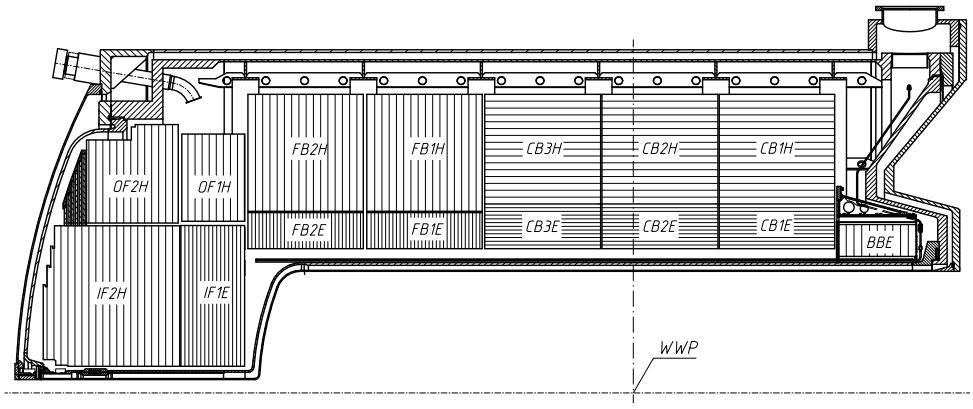

Figure 3.6: Longitudinal cross section of the Liquid Argon calorimeter showing its segmentation into eight wheels.

The LAr is a sampling calorimeter composed of alternating absorber layers and liquid argon filled gaps. The electromagnetic section consists of 2.4 mm thick lead absorber plates separated with 2.35 mm wide liquid argon sampling layers. The liquid argon acts as an active material between the high voltage and readout cells which are mounted on the absorber plates. Shower particles crossing the sampling layer induce a signal by the ionisation of liquid argon atoms. The total depth of the absorber material in

Figure 3.7: Longitudinal cross section of the Liquid Argon calorimeter showing lines of constant radiation length $20X_0$ and constant nuclear interaction length λ

As illustrated in figure 3.6, the orientation of the absorber plates varies between wheels, being horizontal in all three CB wheels and vertical elsewhere.

Both the electromagnetic and the hadronic section are highly segmented in the transverse and longitudinal direction with about 44000 cells in total, which is illustrated in figure 3.8. The cell granularity is finer in the forward direction on account of the higher particle concentration and in the electromagnetic section in order to resolve the compact electromagnetic showers induced by electrons and photons. The longitudinal segmentation varies from three (central) to four (forward) layers of cells in the electromagnetic section, where the first layer has a thickness of three to six radiation lengths, and from four to six layers in the hadronic section. Transversally a basic granularity of the electromagnetic readout cells is twice the Molière radius[9] $2R_M$ measured at the entrance of the electromagnetic calorimeter. In the BBE, CB1 and CB2 the transverse resolution is relatively poorer. For particles incident from the interaction point, the laterally projected

[7]The radiation length is a material dependent constant, which measures the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung. A good approximation of the radiation length X_0 is [99] $X_0 = \frac{716.4\mathrm{g/cm}^2 A}{Z(Z+1)\ln(287/\sqrt{Z})}$, where A is the atomic weight and Z the atomic charge.

[8] Nuclear interaction length λ is defined as an average free path between inelastic interaction.

[9]The Molière radius R_M is defined i.e. in [100, 101] and is a measure of the transverse dispersion of the electromagnetic shower. Roughly 95% of the shower is contained within the radius of $2R_M$.

cell size in the electromagnetic stack ranges between 5×5 cm^2 in the forward and 7×13 cm^2 in the central wheels.

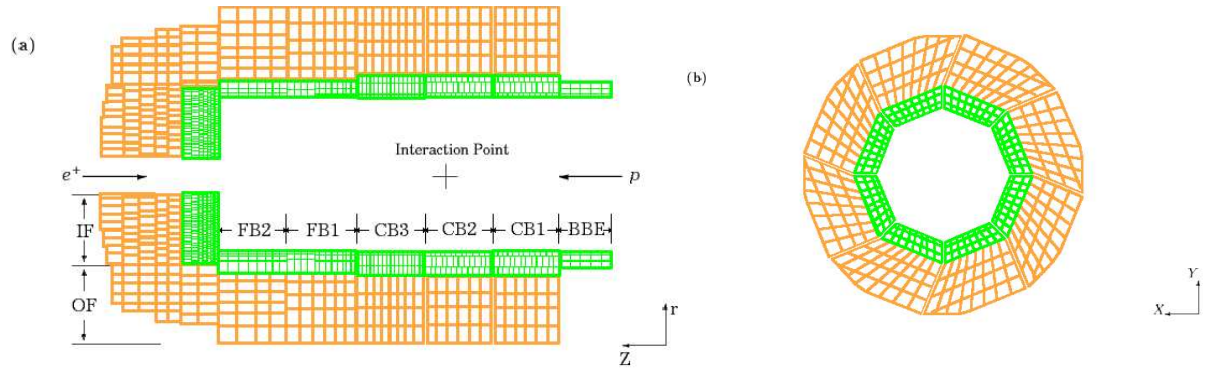

Figure 3.8: Longitudinal (a) and transversal (b) cross section of the Liquid Argon calorimeter illustrating the granulity of the readout system.

The energy resolution of the LAr calorimeter, determined in the test beam measurements [102, 103] is

$$\frac{\sigma^{el}(E)}{E} = \frac{12\%}{\sqrt{E/\mathrm{GeV}}} \pm 1\% \tag{3.4}$$

for the electromagnetic section and

$$\frac{\sigma^{had}(E)}{E} = \frac{50\%}{\sqrt{E/\mathrm{GeV}}} \pm 2\% \tag{3.5}$$

for hadronic showers.

During the reconstruction process, the readout cells of the LAr firstly undergo the noise suppression algorithm and later are assigned to *clusters*, which are contiguous formations of cells that are likely to contain the cascade of a single incident particle.

In the backward part of the H1 detector, the SPACAL [104] is used with the main purpose of detection of the electron scattered under the polar angles $153° < \theta < 177.5°$. This coverage corresponds to the measurements in the kinematic range of $4 < Q^2 < 150\,\mathrm{GeV}^2$. Like the LAr calorimeter, SPACAL is a sampling calorimeter with an inner electromagnetic and outer hadronic section. Both parts are fabricated of long scintillating fibres placed parallel to the beam axis and embedded in the lead absorber material. Charged shower particles in the induced showers are detected by the excitation of molecules in the scintillator material, which trigger light impulses in the fibres. The light is transmitted to photomultiplier tubes at the backward end of the fibres, where the impulses are converted into electrical signals. In this analysis SPACAL is used to veto low Q^2 DIS events.

3.2.3 Luminosity system and electron tagger

The H1 luminosity system (LUMI) [105] make use of the Bethe-Heitler process $ep \to ep\gamma$ predicted in QED with high accuracy. Both electron and photons are mostly scattered

under the small angles and leave the detector through the beam pipe. They are detected by the coincidence of the signal in two calorimeters, the electron tagger (ETAG) and the photon detector (PD). The overview of the system is given in the figure 3.9.

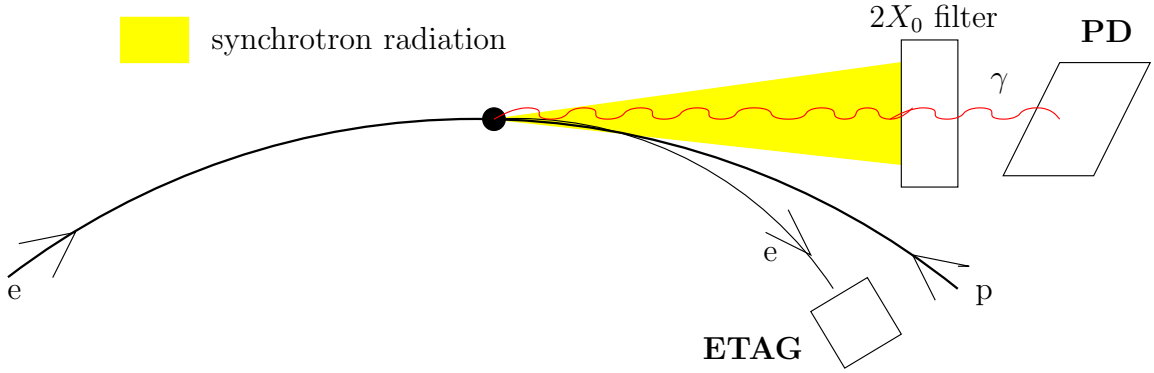

Figure 3.9: The H1 luminosity system. Photons and electrons are detected in the electron tagger (ETAG) and the photon detector (PD).

The PD is located at $z = -103$ m and is a tungsten (absorber plates) / quartz (active material) fibre based sampling calorimeter with the sufficient depth to fully contain electromagnetic shower of about $25X_0$. The position of the PD makes it strongly exposed to the synchrotron radiation. The beryllium based filter with depth of $2X_0$ is placed in front of the main calorimeter, in order to protect it from the synchrotron radiation. Incorporated into the system are two electron taggers, placed in positions $z = -5.7$ m (ETAG6) and $z = -44$ m (ETAG44). The ETAG6 is a SPACAL type calorimeter containing scintillating fibres. The ETAG44 have been installed already during the HERA I running period and is based on scintillating crystals. Both of the electron taggers are used to tag the photoproduction events ($Q^2 < 0.01\,\mathrm{GeV}^2$).

3.2.4 Time-of-flight system

The Time-of-Flight (TOF) system consists of plastic scintillators installed at various places within the H1 experiment. The TOF system delivers precise timing information used to reject non-ep background events which are asynchronous to the strict 96 ns period of *ep* collisions at HERA. In particular two scintillators of the veto-wall (VETO) at $z = 8.1$ m and $z = 6.5$ m reject events with a significant time offset to the HERA clock.

3.2.5 Trigger system and data acquisition

The main objective of the trigger system is to provide a fast decision for the acquisition of interesting *ep* events while sorting out background events. Background sources are mainly due to reaction of the protons e.g. reaction of the protons with the gas in the beam pipe (beam-gas) or with the material of the beam pipe (beam-wall). Beam halo muons and cosmic muons also yield significant amount of background. The background

rates exceed the *ep* interaction rates typically by two orders of magnitude, so the efficient trigger system is needed for successful experimental running.

The H1 trigger system consists of four levels (L1 - L4), of which all but L3 was in operation during HERA II data taking period. The H1 trigger system, consists of several levels, denoted L1 to L4. Only after an event has been accepted by all operating systems, it is written to tape and analysed by the offline reconstruction. An illustration of the data flow through the H1 trigger system is shown in figure 3.10.

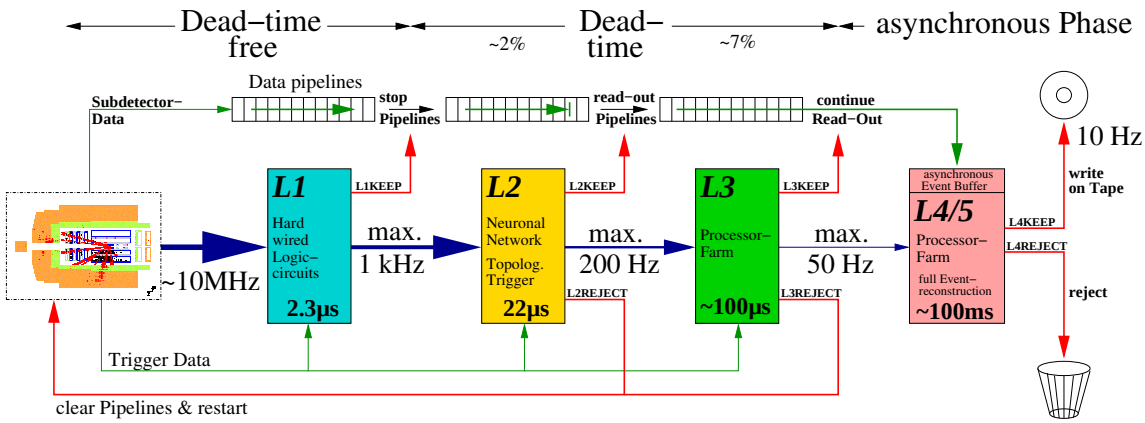

Figure 3.10: The designed data flow through the H1 trigger system.

The first level trigger, L1, consists of around 200 trigger elements (TE) providing fast information from different detector subsystems. The central trigger logic combines TEs into 128 subtriggers, the majority of which are designed to select a variety of physics processes, although some are used to monitor background and trigger efficiencies. An event is kept if at least one of 128 subtriggers (s0-s127) give a positive decision. If a specific subtrigger has a too large rate it is consequently prescaled While the trigger decision is being taken, the readout is stored in the peaplines, which enables the L1 trigger to be free of a deadtime. The L1 decision is taken within 2.3 μs and the output of the L1 trigger system is of the order of 1 kHz.

The L2 system decision is derived within 22 μs from one of two independent hardware systems, a topological trigger (L2TT) and a neutral network trigger (L2NN) which both combine the information of several subsystems. A positive decision on L2 stops the pipepines and the whole event is read out. The readout process takes typically $1-2$ ms during which no further data can be collected and is considered as a detector dead time. The designed output rate of the L2 trigger is predicted as 200 Hz, but since the L3 trigger level was not in operation during the most of the HERA II running period, the L2 actually must decrease the rates down to 50 Hz.

The L4 trigger decision is based on a full though simplified event reconstruction and decided in approximately 100 ms. Basing on calculated so called L4 *statements*, selected events are grouped into L4 *classes*. The physics event classes are selected based on a presence of a hard scale, like sufficiently large Q^2 or transverse momentum, or by the output of dedicated physics finders like 'open charm' or 'open beauty'. Events that can

not be assigned to any physics class are classified as 'soft physics' and strongly prescaled (this effect is important for this analysis, for further details see section 4.3.3).

Events passing all the trigger levels are permanently written to tape in two different data formats, *Production Output Tapes* (POTs) and *Data Summary Tapes* (DSTs). The POTs contain the complete event data including raw and reconstructed information, while on DSTs only a subset of predominantly reconstructed quantities is stored, which is sufficient for most physics analyses. DSTs occupy roughly 10 kB per event of disk space, which is an order of magnitude smaller than the data volume of the POTs. In the offline analysis of the *ep* scattering data, calibration constants to the subdetector measurements, as well as information on the detector alignment, are extracted. After the so-called *reprocessing*, the corrected measurements are written to a further generation of DSTs. In this analysis DST3 data reprocessing generation was used.

Chapter 4

Event reconstruction and preselection

This chapter provides the basic informations about the reconstruction of general event properties as well as criteria requested to obtain basic data selection. Section 4.1 gives an overview of the data on which this analysis is based. Section 4.2 explains the reconstruction of the basic event properties, particularly explains the jet reconstruction algorithm. The triggering issues are explained in details in section 4.3, while additional requirements aimed to reject non-*ep* background are discussed in section 4.4.

4.1 Data overview

The analysis presented in this thesis is based on the data collected by the H1 detector in the years 2004-2007 and represents the whole HERA II data sample. The data taking strategy of the H1 experiment groups the collected events into runs of of approximately 1 hour duration. Depending on the quality of collision environment and the experimental hardware status the runs are classified offline as *good*, *medium* or *poor*. The analysis presented in this thesis does not use runs degraded as *poor*. In addition it is required that the high voltage of all the essential subdetectors is at its nominal value and their readout is functional. The essential subdetectors for this analysis include: CJC1, CJC2, LUMI, LAr calorimeter, SPACAL, CIP, BPC, TOF and VET systems (see section 3.2 for systems description).

Table 4.1 contains the detailed information about used H1 data periods, including the charge of the lepton beam and its helicity. The polarisation of the electrons has not been used in this analysis, but since it is commonly used to define data taking periods, it is included in the table. The total integrated luminosity of all considered HERA run periods amounts to 340.1 pb^{-1}.

In figure 4.1 the event yield for inclusive prompt photon selection (selection criteria will be discussed later and are summarized in section 5.4) is presented. The yield is stable over the time and amounts to roughly 440 events per 12 pb^{-1} of integrated luminosity. Visible

Year	Beams	Helicity	Run Range	$\int \mathcal{L}dt$ [pb^{-1}]
2004	e^+p	+	367257 – 376562	16.7
	e^+p	−	376810 – 386696	23.7
	e^+p	+	387537 – 392214	8.7
	e^-p	+	396674 – 398679	0.16
2005	e^-p	+	399629 – 402634	1.3
	e^-p	−	402993 – 414625	34.1
	e^-p	+	415726 – 427474	30.0
	e^-p	−	427872 – 436893	34.8
2006	e^-p	−	444307 – 458154	35.6
	e^-p	+	458841 – 466997	21.3
	e^+p	+	468531 – 485678	57.6
	e^+p	−	485715 – 486072	1.06
	e^+p	+	486073 – 490161	18.7
	e^+p	−	490162 – 492541	10.3
2007	e^+p	−	492559 – 500611	45.9

Table 4.1: Summary of the data periods used in this analysis.

variations are caused by the changing trigger efficiency and variable beam qualities.

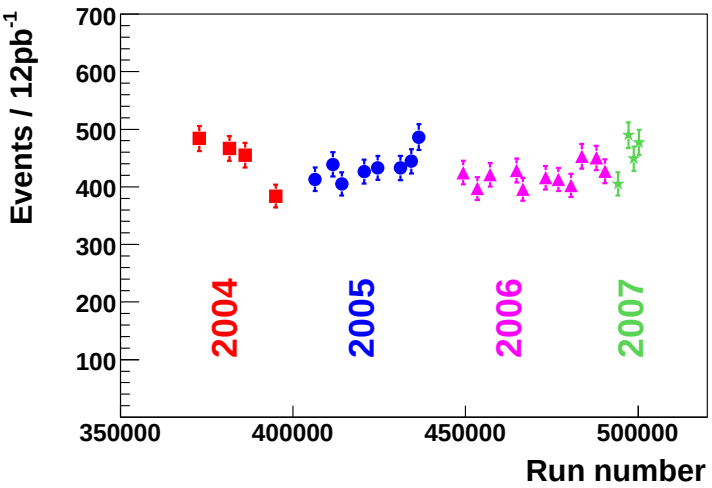

Figure 4.1: Event yield for inclusive prompt photon selection. Each point corresponds to the number of events selected per 12 pb^{-1} of integrated luminosity.

4.2 Event reconstruction

Electronic signals collected directly during the data taking are subject to algorithms that builds more physical objects. H1 track finding strategy (based on the algorithm described in [106]) and calorimetry clustering is in details explained in [67]. For an overview of the primary vertex reconstruction see [107]. On the following pages the scattered electron identification and jet algorithm is presented.

4.2.1 Scattered electron identification

Two scattered electron finding algorithms are used in this analysis, with one looking for electron in LAr calorimeter, the second in SPACAL.

The algorithm of scattered electron finding in the LAr calorimeter is based on the QESCAT finder which main characteristics are described in [108]. The LAr clusters composed of at least three cells, with energy higher than 5 GeV and transverse momentum higher than 3 GeV are selected. Additional cuts on cluster quality criteria are introduced such as electromagnetic energy fraction, cluster compactness, radius and isolation. Finally, all the scattered electron candidates are required to have an associated track with a distance of closest approach below 12 cm.

In SPACAL all the clusters with electromagnetic energy above 5 GeV and radius below 4 cm are selected. The BPC and BST tracks are used to validate the electron candidate position, but are not explicitely required.

From the selection of all LAr and SPACAL electron candidates the one with the highest transverse momentum is selected as the scattered electron and used to veto DIS events as explained in section 5.1.

4.2.2 Jet algorithm

Due to the confinement effect, partons from hard interaction can not be observed directly. Instead, during the hadronisation process many colourless secondary particles are created and leave their signal in the detector. Those final state particles can be grouped back into jets containing the information about the initial partons produced directly in the collision. A good jet algorithm should construct objects which well map the parton configuration and be collinear and infrared safe. Infrared safety means that adding a soft particle into the final state should not change the jet configuration and is a requirement needed to minimise the influence of an electronic noise. Collinear safety means that two parallel particles are exactly equivalent to one particle with the sum of the momenta of the pair. This requirement accounts for resolution effects of the detector. In this analysis the jet k_T algorithm [109] is used which is both collinear and infrared safe. The algorithm works as follows:

1. The initial object list is created from all the particle candidates being all the tracks

and energy deposits in the detector[1]. Special care was taken to avoid double counting.

2. A distance d_i for each object and d_{ij} for every pair of objects is calculated:

$$d_i = P_{T,i}^2, \quad d_{ij} = min(d_i, d_j) \cdot \Delta R_{ij}^2 / R_0, \tag{4.1}$$

where $\Delta R_{ij}^2 = (\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2$ is the distance squared of the two particles in the $\eta\phi$-plane and R_0 is the resolution parameter.

3. d_{min} is defined as the smallest from all d_i and d_{ij}. If d_{min} is taken from d_i set, object i is removed from the collection of objects and becomes a jet. If, on the contrary, d_{min} belongs to d_{ij} set, objects i and j are removed from the objects lists and new combined object is created following given recombination scheme and added to the objects list.

4. The procedure is repeated until all objects are removed from the object list.

In this analysis objects are combined following the P_T-weighted recombination scheme, where transverse momentum P_T, pseudorapidity η and azimuthal angle ϕ of the composite object are given by:

$$P_{T,ij} = P_{T,i} + P_{T,j}, \quad \eta_{ij} = \frac{P_{T,i}\eta_i + P_{T,j}\eta_j}{P_{T,ij}}, \quad \phi_{ij} = \frac{P_{T,i}\phi_i + P_{T,j}\phi_j}{P_{T,ij}}, \tag{4.2}$$

and the resolution parameter is chosen to be $R_0 = 1$ as suggested in [110].

The jet k_T algorithm combines all the final state particles unambiguously into jets such that every particle belongs to one and only one jet.

4.3 Triggering

The H1 trigger strategy has been explained in section 3.2.5. Particularly in this analysis the subtrigger s67 is used. The trigger efficiency is studied using independent monitor triggers and accounted for.

4.3.1 Level 1 trigger

The s67 subtrigger is created with the purpose of a fast detection of the electron scattered under a wide angle. The main component of its definition is the existence of a concentrated electromagnetic energy deposit in LAr calorimeter. It used to be one of the most important subtriggers for the H1 experiment, triggering high Q^2 events and forced, if beam conditions

[1]The same jet algorithm is performed on both hadron and parton levels. In the case of hadron level the initial object list is built out of all the final state particles. For a parton level, all the particles before the hadronisation step are included.

allow stable running, to run without any prescaling. Particularly in this analysis s67 is used to detect photon candidates.

LAr calorimeter trigger works on so called Big Tower (BT) regions chosen such as to point to the nominal vertex as outlined in figure 4.2 for θ structure, additionally divided into 16 or 32 sectors in ϕ. The main condition of s67 subtrigger, *LAr_electron_1* (TE1), detects events with at least one BT with electromagnetic energy deposition above the threshold (of roughly 5 GeV in the central region and up to 25 GeV in the forward). The hadronic energy is ignored. The details of LAr calorimeter trigger setup can be found in [111].

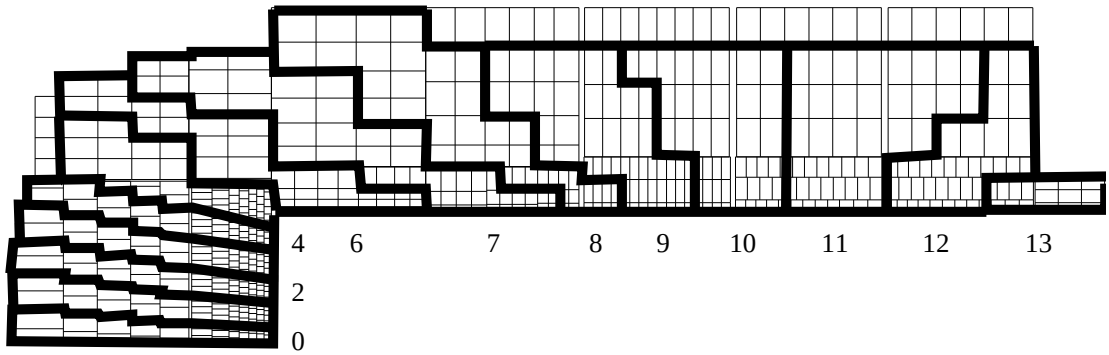

Figure 4.2: The θ distribution of LAr Big Towers.

In order to decrease the probability to trigger on noise in the LAr calorimeter, s67 subtrigger combines TE1 condition with the timing information from the LAr calorimeter and the CIP. Together with the veto on special beam related background signatures provided by the TOF system and the CIP the rate of the subtrigger is decreased to acceptable values. Beginning with the run 483314, the additional veto condition detecting halo muons has been added to the s67 definition. This extra requirement was applied to 20% of the total selected luminosity.

Figure 4.3 presents s67 related rates for an example of 2006/07 e^+p data as a function of time and instantaneous luminosity. The mean TE1 rate (figure 4.3a), being around 20 Hz is decreased by a factor of 10 by the background veto conditions included in the s67 definition, producing a final subtrigger rate on the 2 Hz level (figure 4.3b). The rate is low enough for the L1 trigger to run with a prescale of 1.003 averaged over the whole HERA II running period. The prescale factor for the example run period is plotted in figure 4.3c. Figure 4.3d presents the s67 trigger rate plotted against instantaneous luminosity. The runs taken in the beginning of the fill, characterised by high beam currents produce high instantaneous luminosities. Since in the case of studied trigger, hardly any correlation between subtrigger rates and the instantaneous luminosity is visible, hence one can conclude that s67 triggered events are still mostly non-ep collision background.

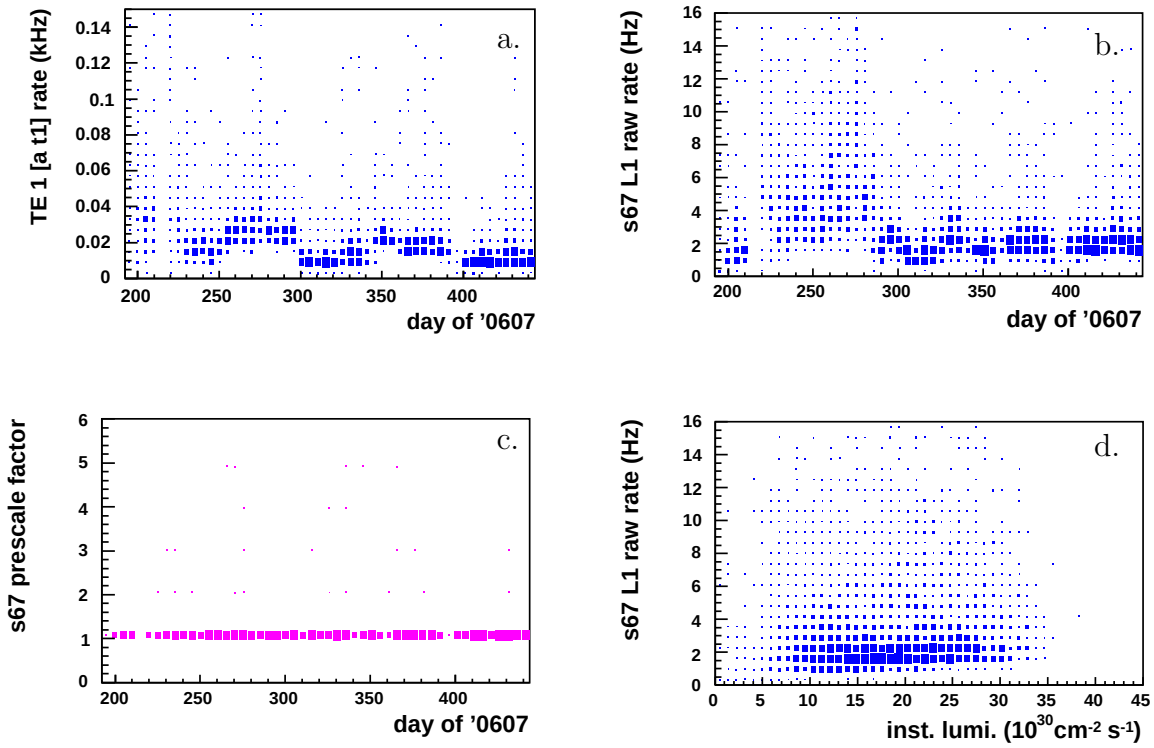

Figure 4.3: Rates for the TE1 (a), s67 (b), s67 prescale factor (c) as a function of time and the s67 trigger rate as a function of instantaneous luminosity (d) for the 2006/07 e^+p data.

4.3.2 Trigger efficiency

Trigger efficiency has been studied using events triggered by a set of independent conditions, such as detection of the scattered electron in the SPACAL or ETAG.

The main trigger monitor sample is triggered by a localised energy deposition (E^{IET}) in the outer region[2] of the SPACAL. Three different subtriggers are used with different energy thresholds and different prescale factors. The subtrigger s3, with the highest energy threshold of 9 GeV has a prescale equal to one, while subtriggers with lower energy thresholds are consequently downscaled. The radial position (R_{Spacal}) of the electron candidate is further verified on L2. The L2 condition for s9 has been changed in 2005 from run 421376 on from $R_{Spacal} > 30$ cm to $R_{Spacal} > 40$ cm.

The alternative monitor sample was used to evaluate possible systematic bias of the trigger efficiency determination. It has been triggered by the scattered electron in the electron tagger (E^{ET6}) placed 6 m away from the main detector and a sufficient number of high

[2]The outer region is defined as everything except the inner region being the square $8.1 \times 8.1\text{cm}^2$ in the centrum of the SPACAL.

energetic tracks in the tracker. Subtriggers s82 and s83 were used. The exact definition for the s82 subtrigger has been changed in the year 2006 from demanding at least 3 tracks with $p_T > 400\,\text{MeV}$ to at least one track with $p_T > 900\,\text{MeV}$. The change has been in charge from run 488289 on. The subtrigger s83 has been additionally verified by L2 topological finder in order to detect D^* or di-jet events and was available until run 487915.

Since the statistics of the alternative monitor sample is insufficient for a reliable determination of the trigger efficiency, the subtrigger s61, based on both SPACAL energy deposition (including the inner part of the calorimeter) and the tracker, has been included.

Table 4.2 summarises the subtriggers used to determine the efficiency of s67. The overlap between two samples amounts to 50% of the main monitor sample.

	Subtrigger	**L1 Condition**	**L2/L3 Condition**
Monitor Sample	s0	$E^{IET}(outer) > 6\,\text{GeV}$	$R_{Spacal} > 20$ cm
	s3	$E^{IET}(outer) > 9\,\text{GeV}$	$R_{Spacal} > 30$ cm
	s9	$E^{IET}(outer) > 2\,\text{GeV}$	$R_{Spacal} > 30(40)$ cm
Alternative Monitor Sample	s82	$E^{ET6} > 3\,\text{GeV}$ $NTrack^{400\text{MeV}} > 2$ $(NTrack^{900\text{MeV}} > 0)$	–
	s83	$E^{ET6} > 3\,\text{GeV}$ $NTrack^{400\text{MeV}} > 2$	D^* or di-jet
	s61	$E^{IET}(all) > 6\,\text{GeV}$ $NTrack^{900\text{MeV}} > 0$	–

Table 4.2: Subtriggers definitions used for monitoring s67 trigger efficiency.

The trigger efficiency has been evaluated for clusters passing the same cluster selection as used later in the main selection with the sole exception of the isolation criteria, which has been skipped altogether. A consistent usage of only highly isolated clusters decreases the statistics of monitor sample overall by 70% and particularly in forward direction by 85% introducing a high error on fitted trigger efficiency. The possible bias due to missing isolation criteria has been studied and the results are compatible within statistical errors.

The LAr trigger suffered during the HERA II running period from various localised hardware problems. In order to restrict the measurement to the well triggered events the trigger efficiency has been studied as a function of the variable zr defined as

$$zr = \begin{cases} z & \text{for } \textbf{CB1}, \textbf{CB2}, \textbf{CB3}, \textbf{FB1}, \textbf{FB2} \text{ LAr wheels} \\ r + 300 \text{ cm} & \text{for } \textbf{IE1} \text{ LAr wheel.} \end{cases} \tag{4.3}$$

where z and r are impact cluster parameters. Since the H1 machine has been subject to some major repair efforts, the trigger efficiency was studied as a function of time. It was also evaluated in 16 ϕ sectors sensitive to the definition of LAr Big Towers. Figure 4.4 presents the example of a zr dependence of trigger efficiency for the 2006 year running period. Shaded LAr calorimeter areas with significantly low trigger efficiency are considered as unreliable and are removed from the selection. The same figure presents the efficiency as produced by the MC simulation. One can note that the main hardware problems were properly simulated but in general the simulation overestimates the s67 trigger efficiency.

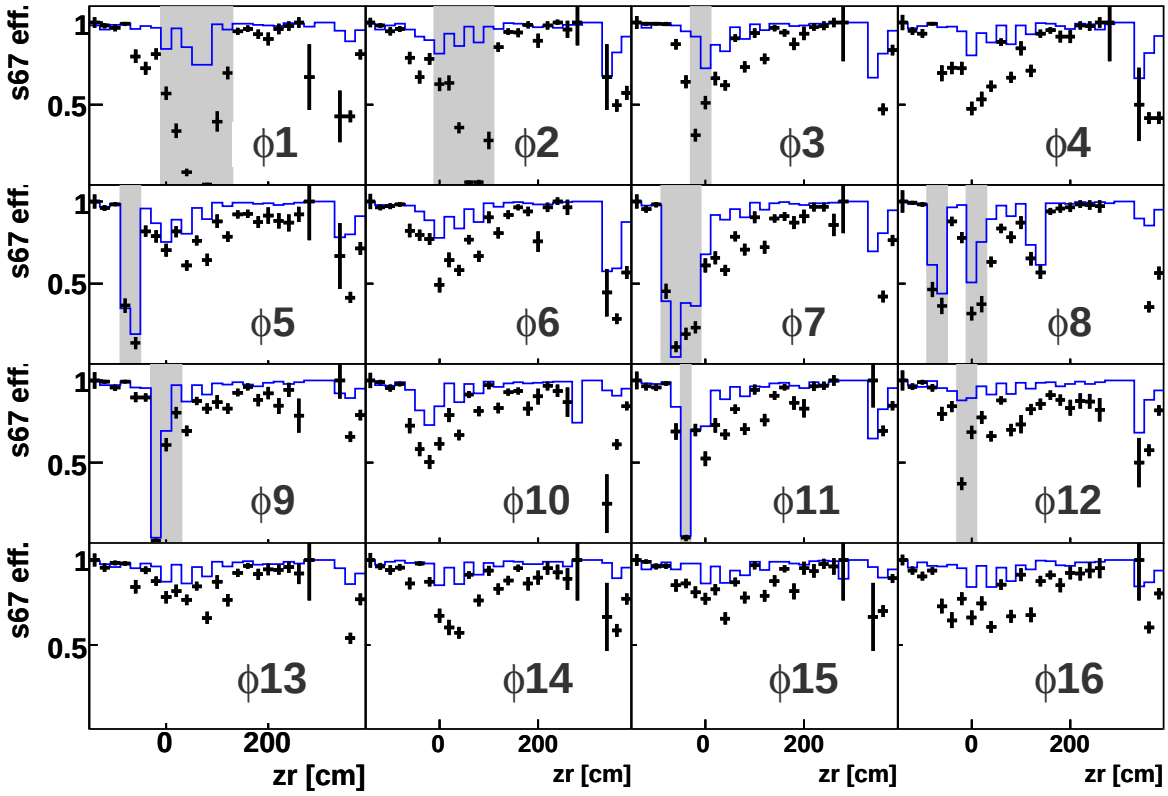

Figure 4.4: The trigger efficiency for 16 ϕ sectors as the function of zr variable. Data points are plotted together with the MC simulation of the trigger efficiency (blue histogram). Shadowed areas indicate regions which are removed from the selection.

The trigger efficiency as a function of the cluster energy has been evaluated for each year and for each LAr calorimeter wheel. An example of determined efficiency for the year 2006 is presented in figure 4.5. The step function has been fitted to the points for both data and MC. One can see again that simulated MC trigger efficiency differs significantly from the measured data efficiency. Therefore, for most H1 analyses, the simulated trigger response is neglected. Instead, the trigger efficiency as determined from data is treated as a separate correction. In this particular analysis though, trigger introduces a bias to the vital variables (cluster shape variables - see section 6.1) and for this reason MC trigger

simulation is used. Additionally efficiency correction is introduced:

$$w^{Trigger}(t, \eta^{\gamma}, E_T^{\gamma}) = \frac{eff_{s67}^{data}}{eff_{s67}^{mc}}, \tag{4.4}$$

where eff_{s67}^{data} is the measured s67 trigger efficiency and eff_{s67}^{mc} is simulated s67 trigger efficiency. The application of the $w^{Trigger}$ variable is given in section 8.2.2.

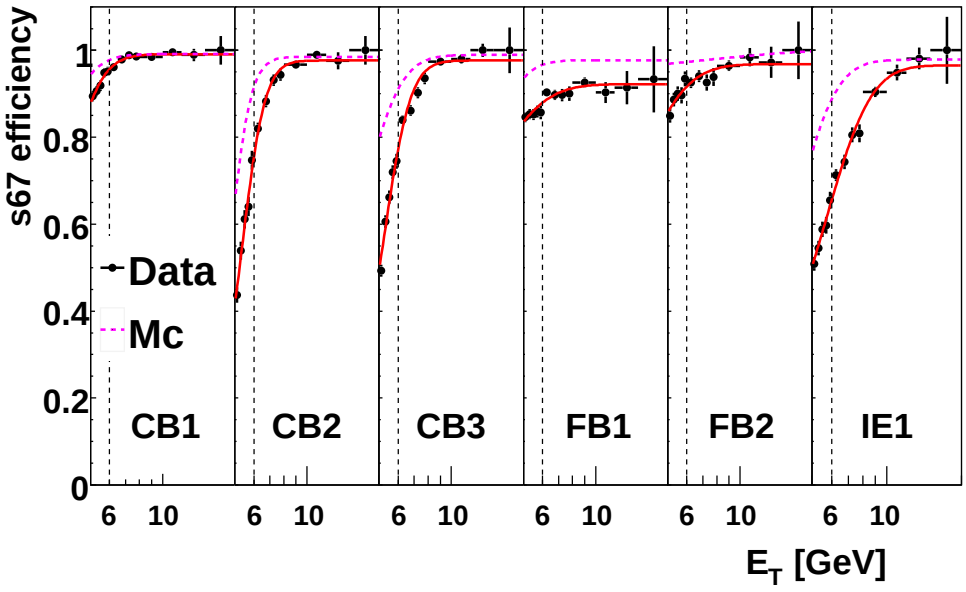

Figure 4.5: S67 trigger efficiency for the year 2006 and for six LAr calorimeter wheels. The measured trigger efficiency is compared to the MC simulation.

4.3.3 Level 4 trigger

Events triggered by s67 subtrigger are subject to event classification on trigger level four (L4)[3] and events not classified as any of the *hard physics* classes are highly prescaled.

The great majority of events triggered by s67 enable L4_HS_ET_CLUST statement on trigger level four. The condition for L4_HS_ET_CLUST is the existence of an electromagnetic cluster with transverse energy larger than 5 GeV. Events with L4_HS_ET_CLUST enabled are classified as class 4 ("High Q^2") events. Figure 4.6 presents the run averaged L4 rejection rate for events triggered by s67 in the example period of 2006/07 e^+p data. The average rejection rate is about 10%. Here, also the additional L1 trigger requirement described above, being in force starting from day $\sim$ 400 of the visualised period may be noted. It greatly improves the quality of s67 triggered events. which is reflected in the drop of the L4 rejection rate.

Trigger level four, is working on fully reconstructed event and thus LAr calorimeter clusters are already formed. The clusters on this level are subject to a rough online energy

[3] L2/L3 trigger level is not used in the main selection for this analysis.

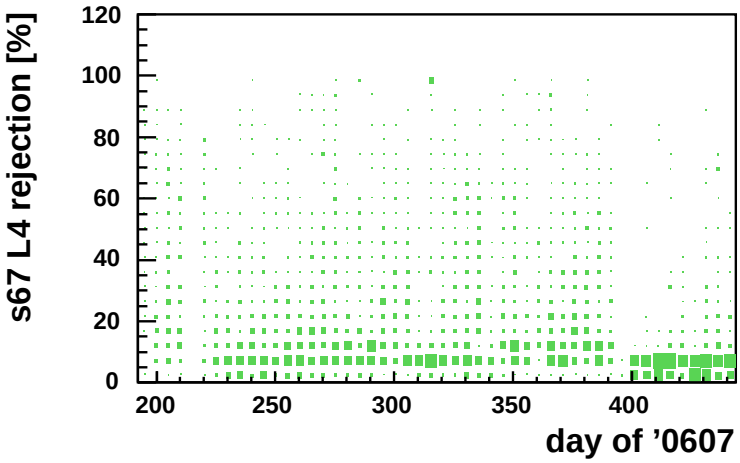

Figure 4.6: Run averaged L4 rejection rate for events triggered by s67 for 2006/07 e^+p data.

calibration, which differs from the precise offline calibration used further in the analysis chain. For this purpose, there might be events passing main selection criteria with higher energy threshold, that did not pass L4 condition and are consequently highly prescaled. Events of this type can not be reliably used in the analysis. For that reason an additional trigger efficiency, including L4, was evaluated. Since event class 4 efficiency is difficult to control (except of highly energetic LAr clusters it may be triggered by a scattered electron in the SPACAL or a general energy deposit in the backward detector region), the efficiency of the L4_HS_ET_CLUST condition, directly corresponding to the measured photon candidate was evaluated. Figure 4.7 shows the trigger efficiency of the combined s67 && L4_HS_ET_CLUST condition with a fitted double step function (motivated by two energy thresholds, one on L1 and one on L4). There is a dramatic drop of the trigger efficiency for low transverse energies, especially in three most forward wheels. Since it is not there for the pure s67 trigger efficiency (figure 4.5), it is attributed to the level four inefficiency with the most probable reason being a significant miscalibration of online cluster energies. Consistently, low E_T clusters are highly prescaled (with the prescale of 40, 60 or 80) preventing a reliable measurement in this region. The final selection is therefore performed for clusters with $E_T > 5.7\,\mathrm{GeV}$ and results are quoted for photons with $E_T > 6.0\,\mathrm{GeV}$. Table 4.3 presents the trigger efficiencies for the year 2006 for few chosen energies.

4.4 Electron-proton interaction event selection

Since the prompt photon production is a rare process, any amount of uncontrolled background results in a significant final error. For this reason extra effort has been taken to obtain a pure ep interaction event selection. The first selection of ep interactions was performed already on the trigger level. Additional offline criteria are discussed in this

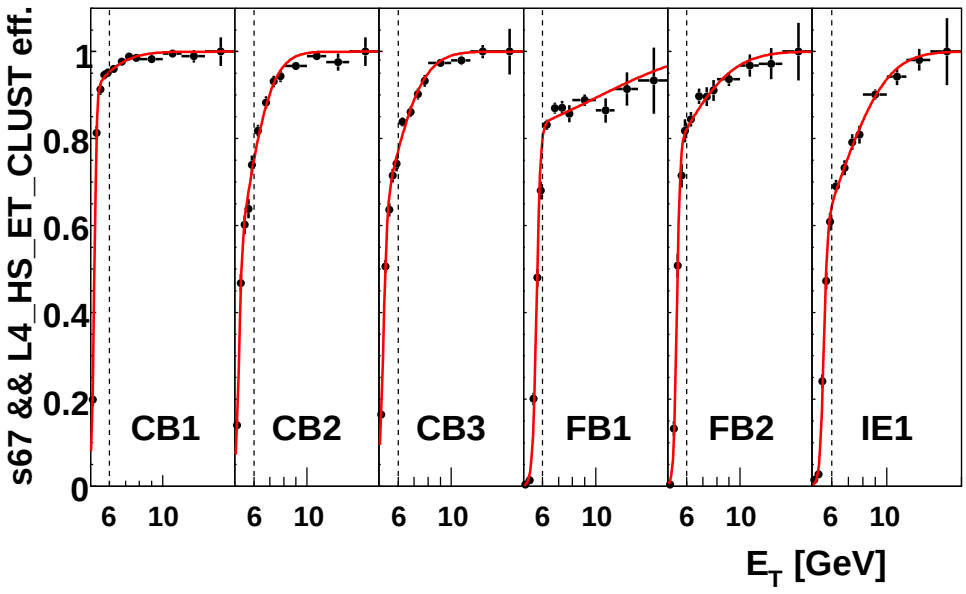

Figure 4.7: S67 && L4_HS_ET_CLUST trigger efficiency for the year 2006 and for six LAr calorimeter wheels.

E_T	CB1		CB2		CB3		FB1		FB2		IE1	
[GeV]	L1	L4	L1	L4	L1	L4	L1	L4	L1	L4	L1	L4
5.0	0.88	0.05	0.42	0.05	0.50	0.07	0.83	0.002	0.86	0.003	0.50	0.002
5.7	0.94	0.94	0.67	0.67	0.70	0.70	0.87	0.48	0.90	0.76	0.61	0.50
6.0	0.96	0.95	0.75	0.75	0.77	0.77	0.88	0.80	0.92	0.82	0.66	0.64
10.0	0.99	0.99	0.98	0.98	0.98	0.98	0.92	0.89	0.97	0.97	0.93	0.93

Table 4.3: Trigger efficiency for 2006 running period for chosen energies. For each LAr calorimeter wheel the value is given for s67 condition (L1) and s67 && L4_HS_ET_CLUST (L4).

section.

The timing information provided by the CJC and the LAr calorimeter is processed offline and events with significant deviation from the nominal value are removed from the sample. For genuine *ep* colision events a primary vertex should be reconstructed nearby the nominal vertex position, so $|z_{vtx}| < 35$ cm cut is introduced. For a correct reconstruction of the primary vertex only events with at least two central tracks are accepted for further investigation.

A set of topological non-ep background finders (summarised in table 4.4) has been applied to the selection. The set consists of 9 out of 10 background finders considered as "safe" (with the sole exception of HALAR finder). Four of them are dedicated to tag beam-halo

events and five to cosmic muon events. Background finders cause the estimated signal selection efficiency drop by roughly 4%, caused mostly by COSLAR finder. Further details can be found in [112, 113].

Finder	**Description**
HALAR	Pure LAr "halo" longitudinal horizontal pattern
HAMULAR	Relaxed LAr "halo" pattern matching a deposit inside the Backward Iron Endcap.
HAMUMU	Horizontal forward muon matching a deposit inside the Packward Iron Endcap.
HAMUIF	Isolated inner forward LAr energy matching a deposit inside the Backward Iron Endcap.
HASPALAR	Isolated inner forward LAr energy matching a deposit inside SPACAL.
COSMUMU	Two opposite muon tracks of comparable direction.
COSMULAR	LAr energy matching the direction of a muon track.
COSTALAR	LAr energy matching the direction of two opposite Tail-Catcher clusters.
COSTRACK	Two CJC tracks with directions exactly opposite in space.
COSLAR	"Long" isolated LAr cluster with a small electromagnetic energy content.

Table 4.4: Definitions of ten topological non-ep background finders considered as "safe".

Chapter 5

Prompt photon selection

Having the data selected as descibed in chapter 4, the prompt photon selection in photoprodiction is performed and described in this chapter. Section 5.1 discusses cuts used to obtain highly pure photoproduction event sample. Prompt photon (section 5.2) and hadronic jet (section 5.3) selection details are given later. The whole selection is summarised in section 5.4.

5.1 Photoproduction event selection

Experimentally, photoproduction events are identified with a scattered electron escaping through the beam pipe. Hence they can be identified by applying a veto on directly measured scattered electron in the SPACAL or LAr calorimeter. This restricts the virtuality to $Q^2 < 4$ GeV for the HERA II setup.

Identification of the scattered electron has been described in section 4.2.1. An electron candidate can simultaneously be identified as a prompt photon candidate, particularly in the forward region of the detector, where no track veto is applied to the photon selection (see section 5.2.2). Since electrons scattered under such wide angle are rare, the majority of those clusters are valid photon candidates (see section 5.2). For this reason, the LAr calorimeter electron veto is applied only if the electron candidate does not fulfil the photon selection criteria.

Since the scattered electron escapes the detection, the event kinematics is calculated using the Jaquet-Blondel method [114]. The inelasticity y is reconstructed as

$$y = \frac{\sum (E - p_z)}{2E_e} \tag{5.1}$$

where the sum runs over energy and longitudinal momentum of all final state particles and E_e denotes energy of the electron beam. The characteristics of quantity $\sum (E - p_z)$ is, that it is by definition insensitive to particle losses in the forward direction where large losses are expected from proton remnant. For events with the electron scattered into the acceptance of the detector the inelasticity tends to be close to one. In order to decrease

the remaining DIS background a cut on high inelasticity is introduced. Extremely forward events which are predominantly due to the interaction of the proton with beam gas are removed with the cut on low values of inelasticity. In this analysis y is restricted to $0.1 < y < 0.7$ range.

The y distribution for selected photon candidates is plotted in fig 5.1a. The data is plotted together with the sum of MC prediction (scaled to the measured cross section), MC signal, MC photoproduction background and MC DIS background. The vertical lines indicate the cut values. The remaining DIS background is estimated to be roughly 2% and its y dependence is plotted in fig 5.1b. The final results are corrected for its presence by subtracting the expected contamination.

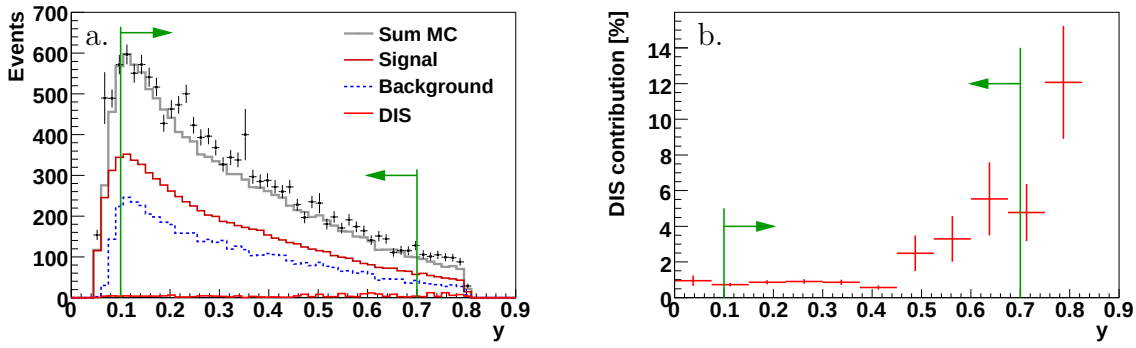

Figure 5.1: Inelasticity estimator y distribution (a) and the estimated DIS events fraction (b).

5.2 Photon selection

Photons are commonly identified in high energy experiments as compact electromagnetic clusters without track pointing to it. A similar approach has been taken in this work with details discussed below.

All LAr calorimeter clusters with transverse energy

$$5.7 < E_T^\gamma < 17\,\text{GeV} \tag{5.2}$$

and in the pseudorapidity range

$$-1.1 < \eta^\gamma < 2.5 \tag{5.3}$$

are used to define the entry point for photon candidate selection. The pseudorapidity range covers almost the whole LAr calorimeter acceptance.

5.2.1 Cluster criteria

Clusters close to LAr calorimeter ϕ and z cracks might possibly suffer from significant energy leakage. Therefore all the clusters with more than 95% of the energy deposited in

cells adjacent to the cracks are removed from the selected sample. In addition the most energetic cell of the cluster is required to be part of the LAr calorimeter wheel in which most of the cluster's energy is deposited. This requirement removes any ambiguity in assigning cluster to particular LAr calorimeter wheel.

Clusters with an electromagnetic energy fraction in the first two layers (E^{γ}_{EM2}) of the calorimeter above a certain threshold are selected. The threshold function has been chosen such as to minimise possible efficiency losses. Figure 5.2 presents the E^{γ}_{EM2} fraction for photon initiated clusters (a) and for hadrons in example of charged pions (b) as a function of polar angle θ^{γ} together with the threshold function. The energy fraction for photons depends on the polar angle due to the difference in angle of impact and thus difference in accessible electromagnetic part of calorimeter. More importantly, it is sensitive to differences of the structure of LAr calorimeter wheels. A significant change is visible around $\theta \sim 50°$ where horizontal absorber planes of CB3 change to vertical absorber planes of FB1. The very low threshold function ($20° < \theta < 30°$) corresponds to FB2 wheel, where the first layer contains G10 material instead of lead. One can note that photons can be found generally above the cut function, while for hadrons the energy fraction in the first two layers in the calorimeter is usually below 5%. The high electromagnetic fractions produced by hadrons for backward detector region ($\sim 140°$) correspond to the edge of hadronic LAr calorimeter.

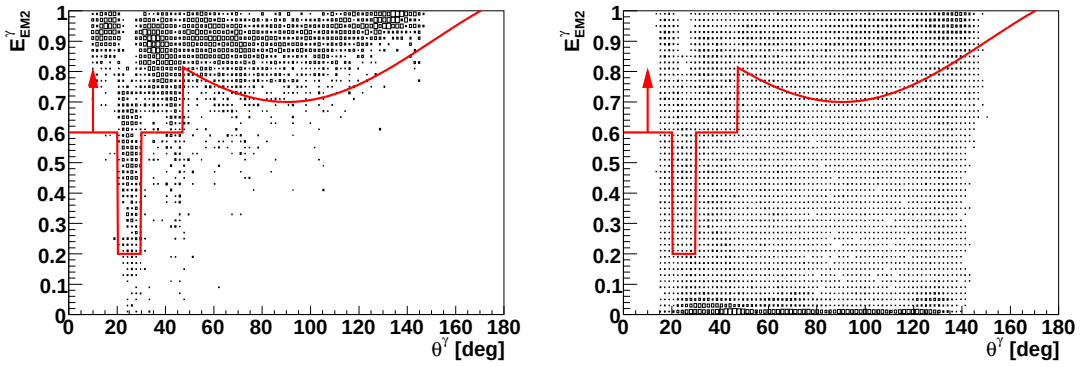

Figure 5.2: The electromagnetic energy fraction in the first two layers of LAr calorimeter of photon clusters (a) and hadron clusters (b) together with the threshold function.

The trigger correction, as discussed in section 4.3.2 is defined by the position and energy of the photon candidate with the assumption that the candidate itself enabled the trigger during the online trigger selection. Since this assumption is not necessary always true, special attention has been paid to ensure its validity. Figure 5.3 presents the difference between candidate cluster and the closest enabled BT (see section 4.3.1) in pseudorapidity and azimuthal angle. One can see the greatest majority of the clusters correspond to an active BT nearby. Interesting observation can be made for a small, but still visible fraction of events where cluster failed to trigger the BT, while the accompanying jet managed to do it. Since jet in most cases balances the photon in azimuthal angle ϕ, those events can be seen close to $\Delta\phi \sim \pi$ as a clearly visible peak two orders of magnitudes lower than the main one. In order to avoid ambiguity of the trigger correction, a cut was introduced on

$\Delta\eta < 0.4$ and azimuthal angle $\Delta\phi < 0.4$ causing the estimated signal selection efficiency drop by 3% on average.

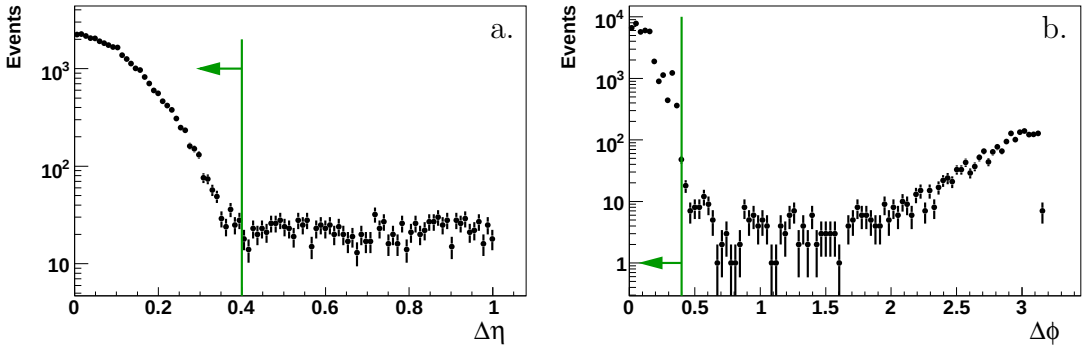

Figure 5.3: The difference between photon candidate cluster and the closest enabled BT in pseudorapidity $\Delta\eta$ (a) and azimuthal angle $\Delta\phi$ (b). The vertical lines indicate the cut values.

5.2.2 Veto on charged particles

Electrons and positrons shower in calorimeter in a very similar way to photons. Contrary to them though they leave signal in all tracking detectors. A tracking veto is thus commonly applied for the photon selection.

In this analysis the main detector providing the veto condition is the CIP. A sensitive area in the CIP is defined for each photon candidate and each CIP layer. A cone of four neighbouring pads in both z and ϕ dimensions around the photon candidate direction is checked for a signal in each of the CIP layers. Candidates with at least one active pad in at least two layers are vetoed as possible charged particles. The efficiency of the veto condition has been studied using both single electrons MC and data and has been found to be above 99%.

Electrons scattered under a wide angle are the main source of charged particle background. Due to the Q^2 dependence of the cross sections, they are expected to appear mainly in the backward region of the calorimeter and in this area the CIP veto has been strengthened by the CJC tracking condition. All clusters in the polar angle range $\theta^\gamma > 45^\circ$ and with distance of closest approach to the nearest CJC track extrapolated to the LAr surface $DCA_{track} < 15$ cm are removed from the sample of photon candidates.

5.2.3 Neutral hadrons background

The main background to the analysis comes from neutral hadrons decaying into multiple photon final states. The great majority of it (87% as estimated in [66]) comes from neutral pions decaying into two photons usually merged into one electromagnetic cluster. Though definite removal of the neutral hadron background is not experimentally possible,

its influence can be reduced. The following section summarises the cluster criteria aiming to reduce the neutral hadron contamination.

Hadrons, produced in the hadronisation process, are naturally expected to be found in jets. Restricting the phase space of the measurement to isolated photons allows background reduction to a reasonable fraction of roughly 50% of the selected sample. Traditionally the isolation criteria is based on a transverse energy deposited around the photon candidate in a cone of one in a (η, ϕ) plane [35,36]. Some more recent results of H1 [43] and ZEUS [42], as well as this analysis follow a z-based approach which is both collinear and infra-red safe.

Given that jet algorithm, as discussed in section 4.2.2, unambiguously assigns a photon candidate to a particular jet, named hereafter as $\gamma - jet$, a variable z (already used in section 1.5) is defined as

$$z = \frac{E_T^{\gamma}}{E_T^{\gamma - jet}}. \tag{5.4}$$

The cut $z > 0.9$ selects photon candidates with very low hadronic activity around it and thus removes a majority of neutral hadron background. Figure 5.4a shows the distribution of the z variable. One can observe that the prompt photon signal indeed produces highly isolated clusters and at $z = 0.9$ the signal distribution drops by almost two orders of magnitude. The high peak at $z = 1$ is caused by events with the photon candidate being the only object within the $\gamma - jet$. The z distribution of the background is relatively flat except of the very highest region.

Background clusters, typically being the result of more than one photon, tend to be wider than clusters initiated by a single photon. A cut on a cluster radius has been introduced. Figure 5.4b shows the radius distribution of all selected clusters. The requirement $R_T < 6$ cm, which is applied to all photon candidate clusters, removes clusters being almost purely background.

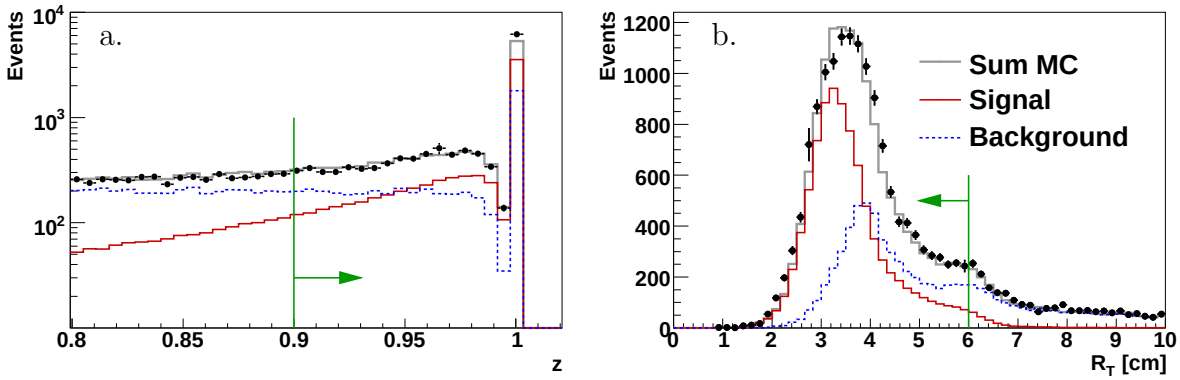

Figure 5.4: The distribution of the z variable describing the isolation of the photon candidate (a) and the transverse radius R_T distribution of the photon candidates (b). Data points are plotted together with the MC prediction. The vertical lines indicate the cut value.

In order to remove background from photons coming from decay of neutral pions and

reconstructed in two separate clusters the invariant mass $M_{\gamma\gamma}$ of the photon candidate combined with the closest electromagnetic cluster having energy above 80 MeV has been introduced. The cut $M_{\gamma\gamma} > 0.3\,\text{GeV}$ removes as well 13% of the signal, mostly photons that converted into electron-positron pair before reaching LAr surface and consistently shower in two separate clusters.

All the cuts described above decrease the contamination of neutral hadrons background to the level of roughly 50%. Determination of final prompt photon content was done on a statistical basis by cluster shower shape analysis (see chapter 6). For a cluster to have well defined shape, valid for the shape analysis, it is required that not all the cells are positioned along one axis. This unlikely configuration happens for 0.7% of all photon candidates, which are removed from the selection.

57 events, possesing more than one valid photon candidate, which amounts to 0.4% of all the selected events, are also removed from the final sample.

All the criterias described above applied to the whole HERA II data sample select roughly 15000 prompt photon candidates. The pseudorapidity η^γ and transverse energy E_T^γ distributions are presented in figure 5.5. The sum of signal and background MC samples descibes the data points reasonably well.

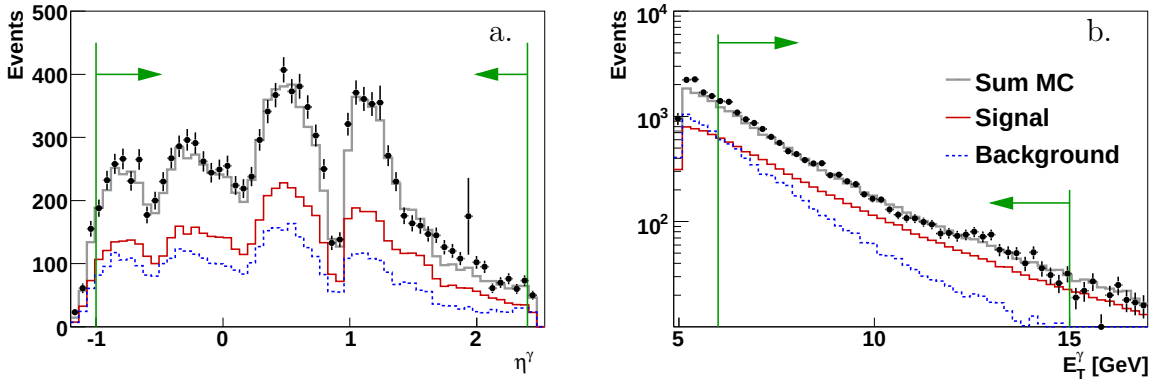

Figure 5.5: Distributions of pseudorapidity η^γ (a) and transverse energy E_T^γ (b) of the selected prompt photon candidates plotted together with the MC prediction. The vertical lines indicate the cut values.

5.3 Exclusive selection and its subselections

A subsample of the inclusive prompt photon selection with additional jet requirement was used for the determination of cross sections in the exclusive phase space (see section 1.5).

Events with the photon candidate fulfilling all the photon selection criteria and with an accompanying jet in the pseudorapidity range $-1.3 < \eta^{\text{jet}} < 2.3$ and transverse energy $E_T^{\text{jet}} > 4.5\,\text{GeV}$ are used to form the exclusive photon selection sample. The cuts are applied on the hadronic jet with the highest energy.

The η^{jet} and E_T^{jet} distributions together with MC predictions are presented in figure 5.6.

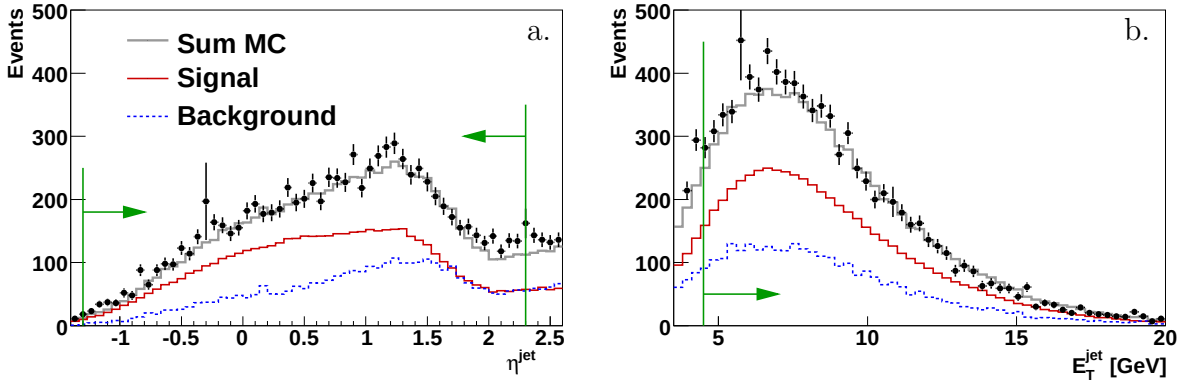

Figure 5.6: The distribution of the pseudorapidity η^{jet} (a) and transverse energy E_T^{jet} (b) of the leading hadronic jet plotted together with the MC prediction. The vertical lines indicate the cut values.

In addition, the exclusive sample was divided into two subsamples based on the x_γ^{LO} variable defined with equation 1.41. The estimator of x_γ is used to divide the phase space into parts with enhanced direct contribution ($x_\gamma^{LO} > 0.8$) and enhanced resolved contribution ($x_\gamma^{LO} < 0.8$). Figure 5.7 presents the x_γ^{LO} and parallel x_p^{LO} distributions together with MC predictions and the cut value separating the two subsamples. One can estimate that roughly 80% of all the signal events in the direct enhanced sample actually comes from genuinely direct events. Resolved and radiated events contribute to the resolved enhanced sample on the level of 53% and 30% respectively.

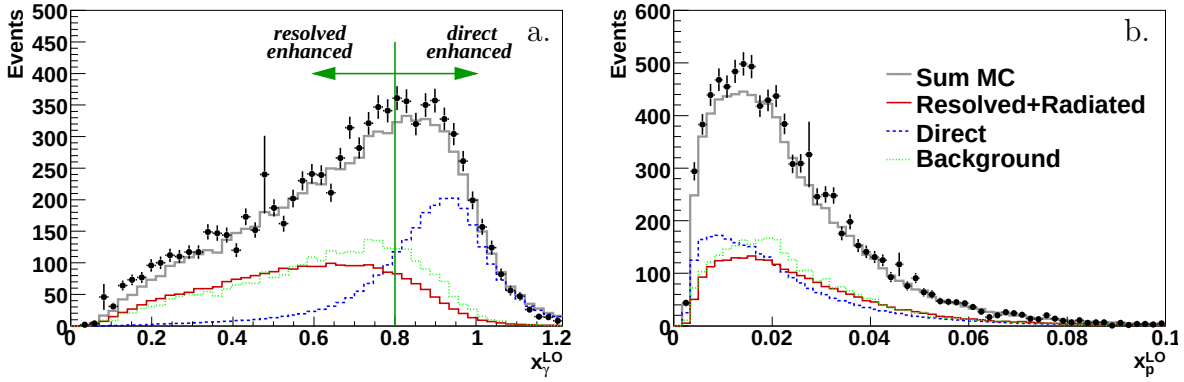

Figure 5.7: The distribution of the x_γ^{LO} (a) and x_p^{LO} (b) estimators. The data points are plotter together with the MC prediction. The vertical line in the x_γ^{LO} distribution indicates the chosen cut value separating direct and resolved enhanced phase spaces.

The effect to be studied in direct and resolved enhanced subsamples is the correlation between hadronic jet and a photon candidate. Two variables defined in the section 1.5, $\Delta\Phi$ and $p_\perp$ are studied and their distributions are presented in figure 5.8.

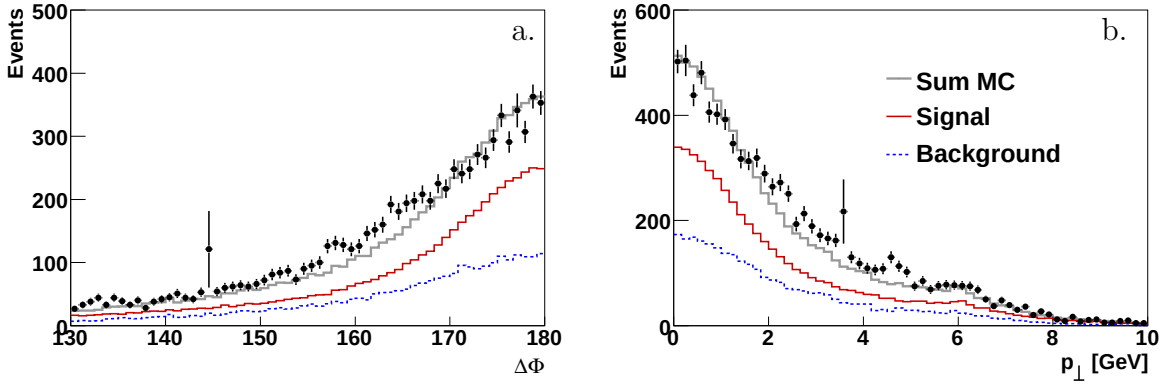

Figure 5.8: The distribution of the $\Delta\Phi$ (a) and $p_\perp$ (b) variables. Data points are plotted together with MC predictions.

5.4 Selection Summary

Table 5.1 summarises all the selection criteria. The inclusive prompt photon selection requires all selection criteria except the cuts on the accompanying jet. The jet cuts are applied for the exclusive selection. The selected exclusive sample is divided into direct and resolved enhanced subsamples based on the x_γ^{LO} cut.

All the cuts were grouped into sets (labeled **A**-**I** in the summary table). The table contains the number of data events and overall signal selection efficiency after various sets of cuts. One should note that the cut on the inelasticity y is introduced twice, once in the ep selection section (**B**) and the second time as a photoproduction selection cut (**C**). The first, weaker, cut was introduced for technical reason in order to decrease the amount of processed data already at the preselection step.

Figure 5.9 presents the signal selection efficiency after various sets of cuts as a function of the transverse energy and the polar angle of the generated photon. For an efficiency dependence of E_T^γ, the transverse energy generator level cut was skipped in order to study the behaviour in a broader energy range. One should note the significant drop of the selection efficiency for low transverse energies already on trigger level (cut set **A**), due to the energy treshold of the trigger condition. The trigger efficiency visible in this plot is lower than discussed in section 4.3.2 (effect particularly visible for high transverse energies) because low efficiency LAr regions which were already removed in the previous discussion. Here, LAr calorimeter fiducial cut selection enter on the level of cut set **E**. The irregular θ dependence of the cut set **A** efficiency resembles the zr trigger behaviour studied previously.

The photon kinematics selection (cut set **D**) restricts the measurement to the interesting phase space. The cluster quality criteria (cut set **E**) cause a significant drop in the overall selection efficiency but are important for selecting reliable candidates with well reconstructed properties. The six wheel LAr calorimeter structure is visible in the θ dependence of the efficiency due to the cuts on LAr z cracks. The hadron suppression cuts (cut set **G**) cause an additional 10% efficiency drop, mostly in the forward region

of the detector, where the $M_{\gamma\gamma}$ cut removes photons that convert into electron-positron pairs already between the interaction point and the LAr surface. Due to the high density of the dead material in the forward region the conversion rate before the LAr calorimeter surface reaches 40%. Photons reconstructed in two separate clusters can not be reliably used in this analysis.

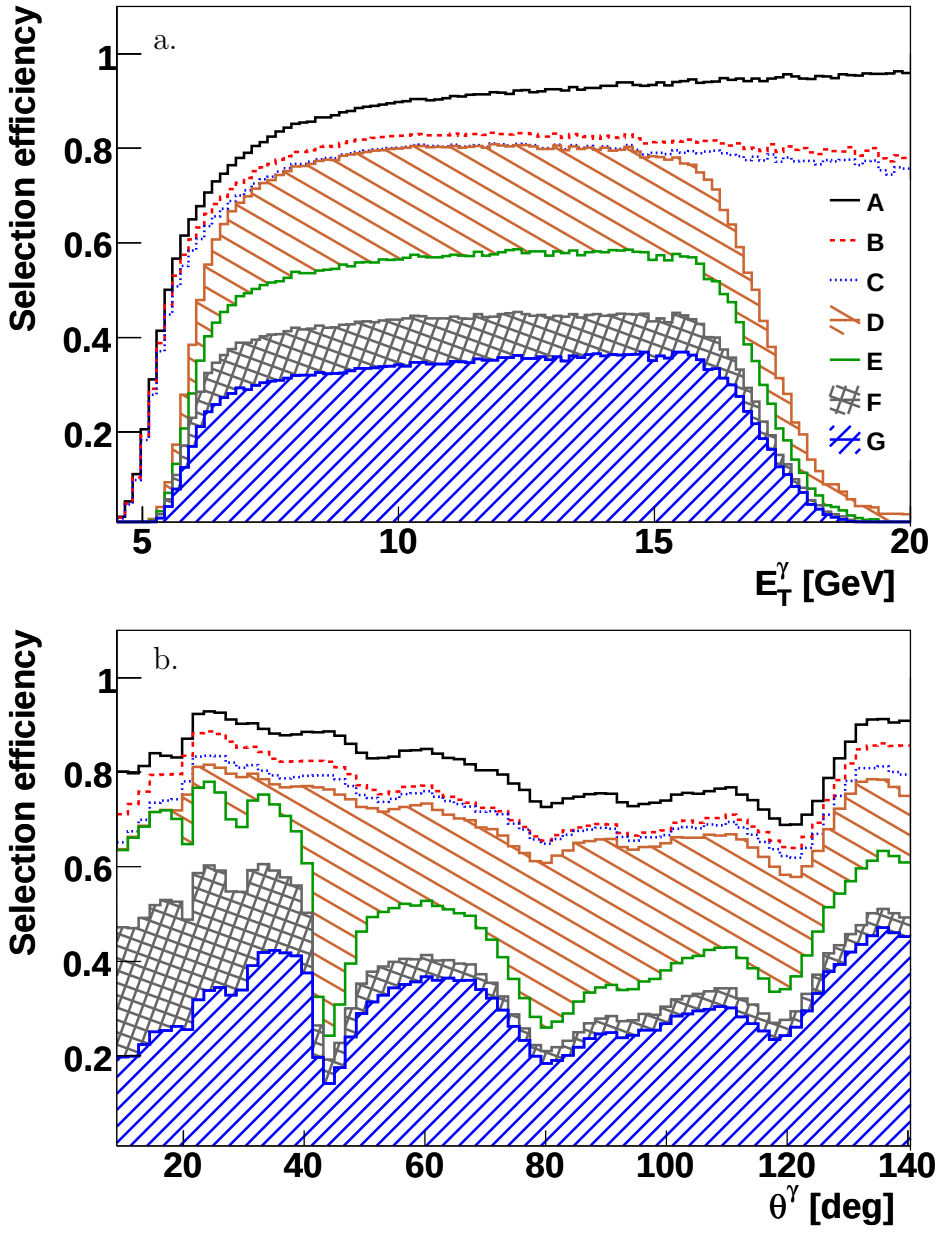

Figure 5.9: Inclusive prompt photon selection efficiency as a function of E_T^γ (a) and η^γ (b) for different cut sets defined in table 5.1.

Selection			Cut	N data	Signal eff.
Direct (Resolved) Enhanced Selection	Exclusive Selection	Inclusive Selection	**HERA II data**	550×10^6	100%
			Online selection (A) s67 subtrigger	65×10^6	82%
			Electron-proton interaction selection (B) Run Quality *medium* or *good* Essential subdetectors HV LAr calorimeter and CJC timing $\lvert z_{vtx}\rvert < 35$ cm Topological background finders veto $0.05 < y < 0.85$ $N^{cent}_{tracks} > 1$	2.6×10^6	76%
			Photoproduction selection (C) SPACAL electron veto LAr calorimeter electron veto $0.1 < y < 0.7$	2.0×10^6	73%
			Photon kinematics selection (D) $5.7 < E^{\gamma}_T < 17\,\text{GeV}$ $-1.1 < \eta^{\gamma} < 2.5$	625×10^3	70%
			Cluster quality criteria (E) $E^{\gamma}_{EM2} > f(\theta)$ LAr calorimeter fiducial cuts LAr calorimeter ϕ and z cracks proximity Wheel unambiguity BT matching	420×10^3	50%
			Charged particles veto criteria (F) $N^{CIP}_{layers} < 2$ $DCA_{track} > 15$ cm (for $\theta^{\gamma} > 45^{\circ}$)	67×10^3	39%
			Neutral hadron suppresion criteria (G) $z > 0.9$ $R_T < 6$ cm $M_{\gamma\gamma} > 0.3\,\text{MeV}$ Cells spread Only one photon candidate	15043	30%
		Accompanying jet selection (H) $E^{\text{jet}}_T > 4.5\,\text{GeV}$ $-1.3 < \eta^{\text{jet}} < 2.3$		8836	
	Event class selection (I) $x^{LO}_{\gamma} > 0.8$ $(x^{LO}_{\gamma} < 0.8)$			3058 (5778)	

Table 5.1: Prompt photon selection summary.

Chapter 6

Photon signal extraction

As already mentioned in chapter 5, the estimated background contamination of the final selection sample amounts to roughly 50%. This chapter provides the detailed information about the method used to extract the precise photon signal fraction in the selected data sample.

As seen in other prompt photon analyses (i.e. [35,43]) the leading background arises from neutral hadrons decaying into multiple photon final states. The study of [66] details the most influential hadrons found in the final background selection to be π^0 (86.8%), η (4.8%) and ρ (3.3%). The mentioned analysis is using single particle MC for a determination of the photon content fraction in final sample and the quoted fraction was used to properly mix the different contributing particles.

6.1 Shower shape variables

The extraction of the photon signal exploits the difference in calorimeter cluster shapes initialised by a single photon as compared to shapes produced by a multi-photon state combined into one cluster. The expected difference is illustrated in figure 6.1 where the sketch of electromagnetic shower initialised by a single photon is compared to the shower produced by two photons originating from π^0 decay. Background clusters are expected to be statistically wider, less compact and more asymmetric. Furthermore, it is likely to start closer to the calorimeter surface, as the probability of conversion increases with the number of initial photons. The expected differences are quantified with six shower shape variables (developed in [43]) which are shortly discussed below.

In the description of transverse cluster shape variables (transverse radius, transverse symmetry and transverse kurtosis), the transverse plane is defined as a plane perpendicular to the momentum vector of the incoming particle, or more practically, to the direction of the vector connecting the reconstructed vertex with the barycentre of the studied cluster. For definitions given below the variable μ_k being the k-th central moment of the transverse cell distribution is used:

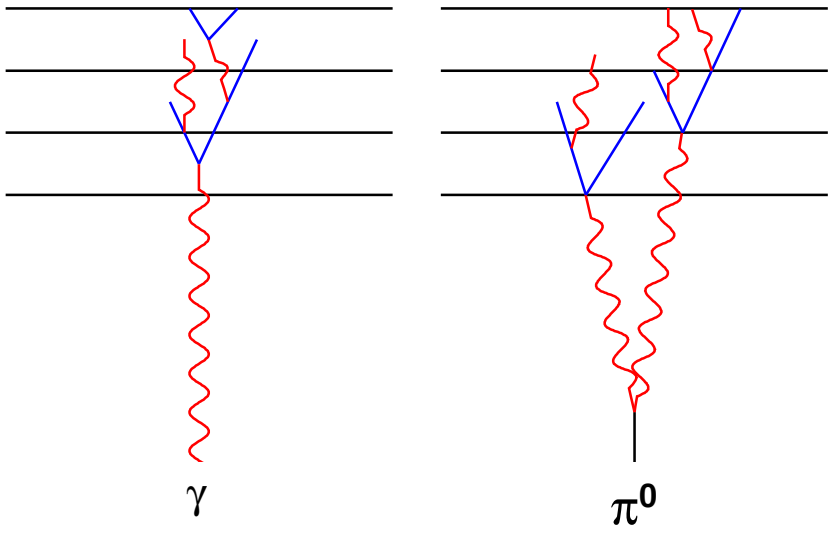

Figure 6.1: Sketch of a calorimeter shower initialised by a single photon and a double photon state originating from decay of π^0

$$\mu_k = \langle |\vec{r}_T - \langle \vec{r}_T \rangle|^k \rangle \tag{6.1}$$

where $\vec{r}_T$ is the transverse component of the cell vector and $\langle \vec{r}_T \rangle = \Sigma_i E_i \vec{r}_{T,i} / \Sigma_i E_i$ is the energy weighted average of the cell position with E_i being the energy of the i-th cell. The studied cluster shape variables are as follows:

- **Transverse radius R_T**
  As seen in figure 6.1 the clusters originating from multiple photon states are expected to be relatively wider. The transverse radius R_T, defined as

  $$R_T = \sqrt{\mu_2}, \tag{6.2}$$

  measures the transverse size of the cluster.

- **Transverse kurtosis K_T**
  Kurtosis is defined as

  $$K_T = \frac{\mu_4}{(\mu_2)^2} - 3 \tag{6.3}$$

  and it measures peakedness of the transverse distribution. Leptokurtic distributions ($K_T > 0$) are the ones with a highly pronounced peak while platykurtic distributions ($K_T < 0$) on opposite are flat-topped. $K_T = 0$ is achieved for a Gaussian distribution [115]. Clusters originating from a single photon are expected to have a more pronounced peak in transverse energy distribution, measured by higher kurtosis values as compared to background clusters.

- **Transverse symmetry S_T**
  Transverse symmetry S_T is defined as the ratio of the spread of the transverse cells distribution along the two principal axes. The eigenvectors of the 3×3 matrix

$$S_{kl} = \langle (r_k - \langle r_k \rangle)(r_l - \langle r_l \rangle) \rangle, \tag{6.4}$$

  where $r_{k,l}$ are the Cartesian coordinates of the transverse cell vector $\vec{r}_T$ constitute the principal axes of the transverse cell distribution. Since the transverse distribution is flat, one of the eigenvalues vanishes and the two remaining ones λ_1 and λ_2 ($\lambda_1 > \lambda_2$) are equal to square radii along the two principal axes. The symmetry is then defined as

$$S_T = \sqrt{\frac{\lambda_2}{\lambda_1}}. \tag{6.5}$$

  Clusters originating from the single photon are expected to be more symmetric than background clusters and thus yield S_T values closer to unity.

- **Hottest cell fraction HCellF**
  $HCellF$ is the cluster energy fraction deposited in the most energetic cell. Since background clusters are initiated by more than one photon, they are expected to be less compact which translates into lower $HCellF$ value.

- **Hot core fraction HCF**
  Hot core is defined as four to twelve (depending on the granularity of calorimeter) contiguous cells in the first two layers around the hottest cell. HCF is the energy fraction deposited in the hot core of the cluster, and being similar to $HCellF$, it is expected to yield higher values for signal-like clusters.

- **First layer fraction FLF**
  FLF variable measures the cluster energy fraction deposited in the first layer of the calorimeter. Since the conversion probability in the LAr calorimeter absorber layer is the same for each photon entering the layer, the variable is expected to be sensitive to the number of initial photons and it is on average higher for multi-photon initiated background clusters.

Distribution of all six shower shape variables used in this analysis is shown in figure 6.2. Data is plotted together with the sum of MC prediction scaled according to the cross section measurement. MC describes the data within the systematic error, which is crucial since the final results determination method relies heavily on it. All the variables yield some signal to background discriminating power and all the effects described above can be seen.

As calculated in [66], the most probable distance between two photons originating from the decay of neutral pion depends on both energy and polar angle of the initial particle. Therefore all size-related shower shapes (all but FLF) are expected to manifest similar dependence for a background sample. In addition, the change of a granulity between LAr calorimeter bins may yield drastic differences in shape of variable distribution, especially if

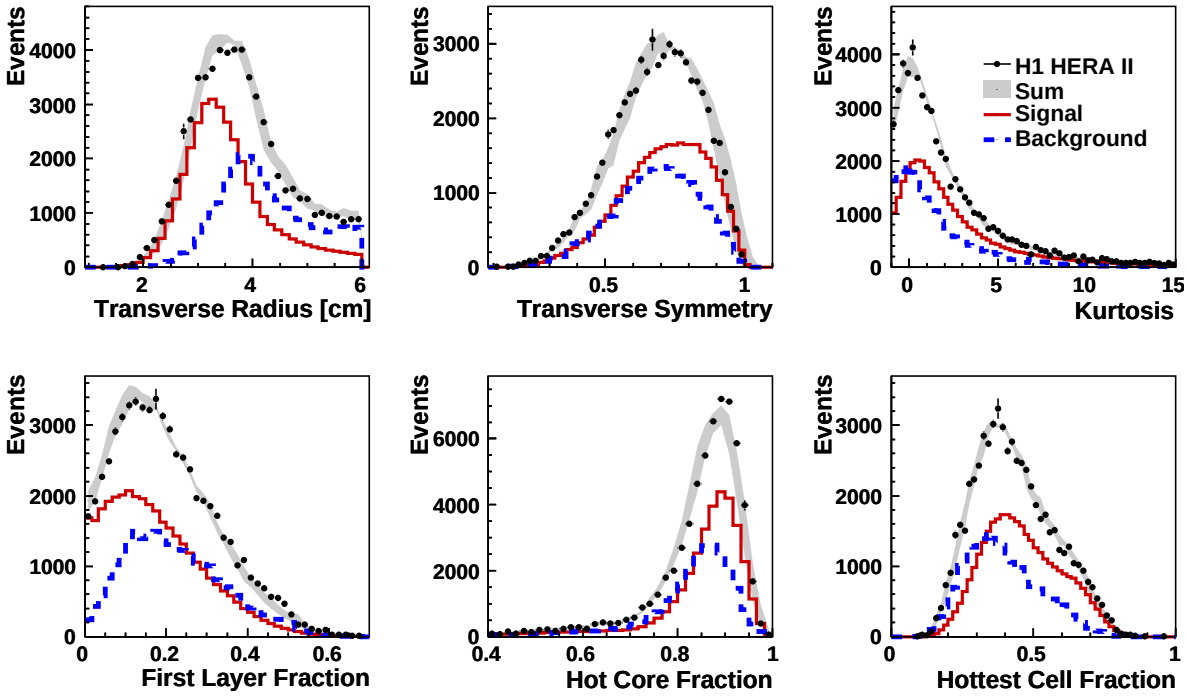

Figure 6.2: Shower shape variables used to discriminate between signal and background clusters. Data is compared to the sum of MC predictions (shadowed gray) as well as signal (solid red) and background (dashed blue) distributions.

that variable has the scale comparable to the granularity itself[1]. Taking into account both effects, all variables have been studied in bins of transverse energy E_T and pseudorapidity η, where pseudorapidity bins correspond to wheels of LAr calorimeter. In figures 6.3 and 6.4 this study is presented for two of the shower shapes: R_T and FLF. The binning used for these plots is consistently used in the following multivariate analysis (MVA) throughout all this chapter and is summarised in appendix table A-1.

Studying the transverse radius distributions, presented in the figure 6.3 one can note the influence of the wheel granulity. Particularly, the R_T distribution is wider in the **CB1** and **CB2** wheels, where LAr calorimeter granularity is the poorest. One can observe the weaking of the discriminating power with the rise of energy of studied particles. Similar effect can be noted for FLF distributions (figure 6.4), which variations in η is reflected in its sensitivity to the technical details of the LAr calorimeter geometry. Particularly **FB2** wheel, with different absorber material in the first layer produces lower FLF values in comparison to other wheels. Since the photon conversion probability has only a weak energy dependence, the shapes of FLF do not vary between transverse energy bins.

[1] As indeed is the case in presented analysis.

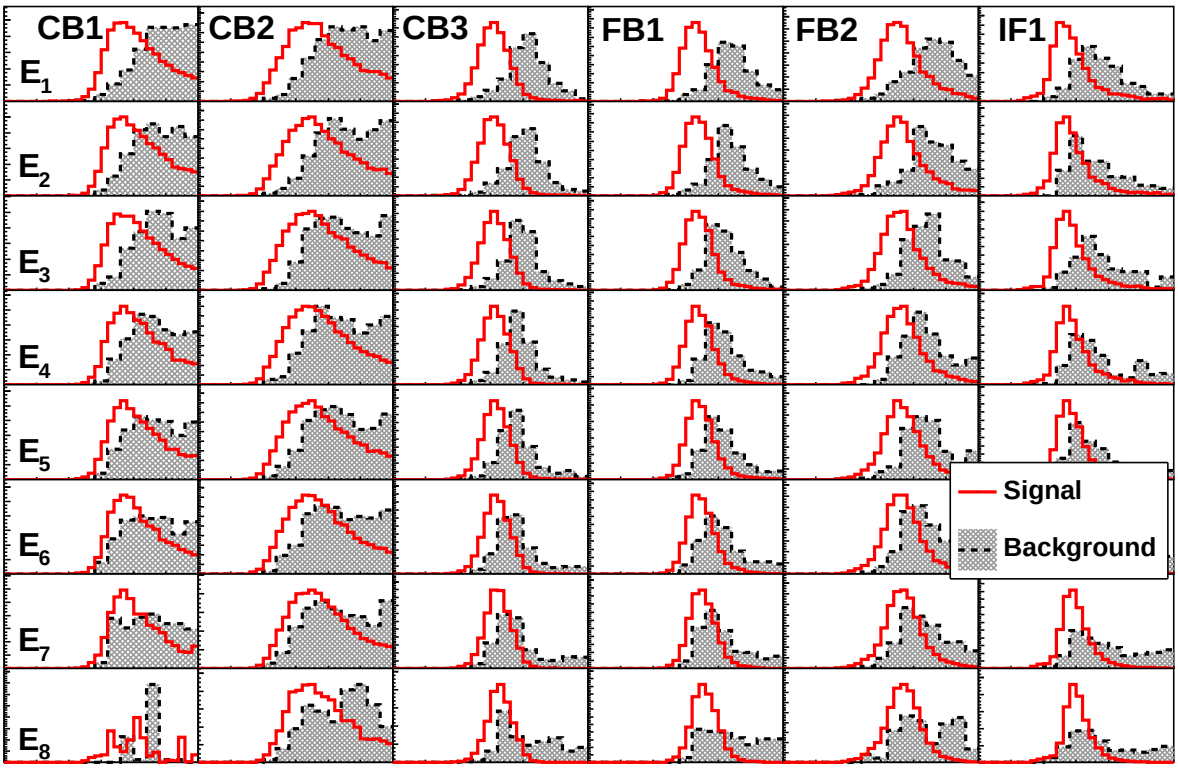

Figure 6.3: Transverse radius signal and background distributions binned in transverse energy and wheel containment of the incoming particle.

6.1.1 Correlations

The high correlations between variables introduced in the previous section may indicate the existence of the redundancy. Figure 6.5 shows correlations between all six shower shape variables studied with signal MC sample (upper right part) and background MC sample (lower left part). No strong correlation is observed.

Further quantification of the correlation may be achieved by calculation of the correlation coefficient ρ. The covariance σ_{uv} between variables u and v and variance σ_u of variable u are given by

$$\sigma_{uv} = \langle u - \langle u \rangle \rangle \langle v - \langle v \rangle \rangle, \tag{6.6}$$

$$\sigma_u^2 = \sigma_{uu} = \langle u - \langle u \rangle \rangle^2. \tag{6.7}$$

The correlation coefficient ρ_{uv} between variable u and v is then defined as:

$$\rho_{uv} = \frac{\sigma_{uv}}{\sigma_u \sigma_v} \tag{6.8}$$

and measures size of the linear dependence between them. Since the correlation does not necessary be of linear type, correlation coefficient needs to be treated carefully, still, may be useful.

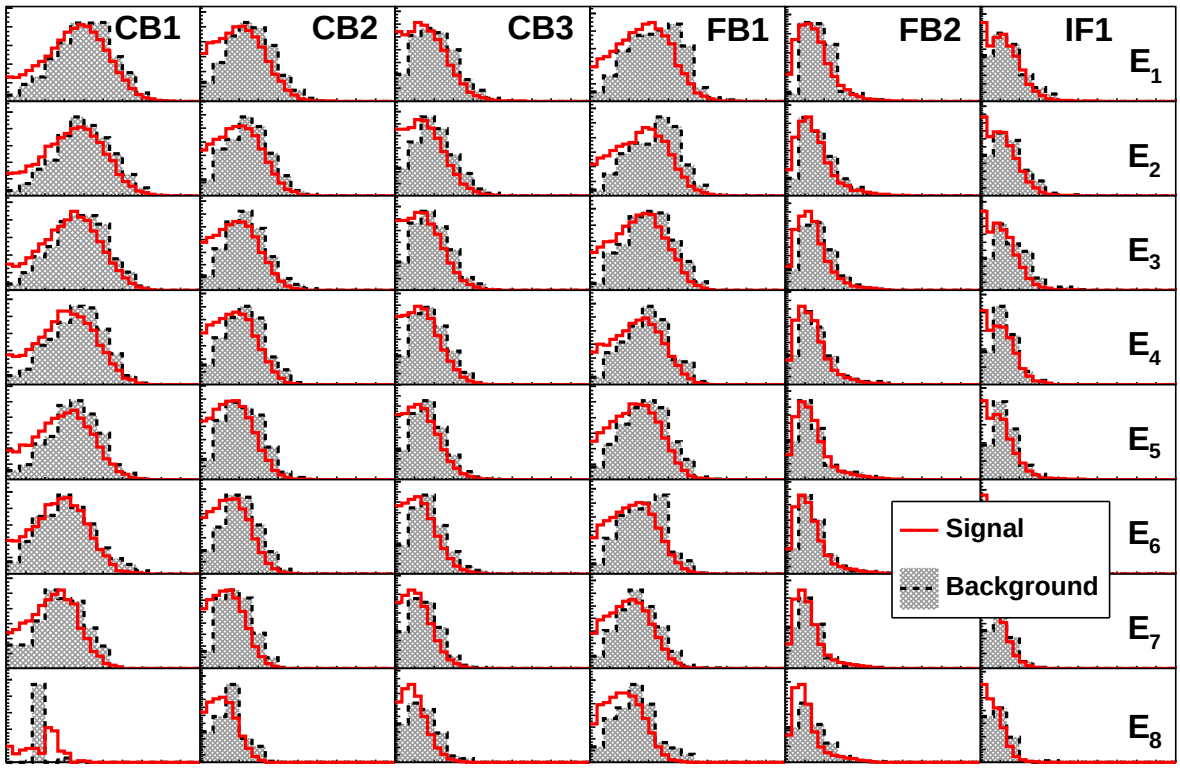

Figure 6.4: First layer fraction signal and background distributions binned in transverse energy and wheel containment of the incoming particle.

Calculated correlation coefficients between all six shower shape variables for signal and background MC samples are given in table 6.1 with the highest value being +0.48 for a correlation coefficient between $HCellF$ and K_T.

	R_T	K_T	S_T	$HCellF$	HCF	FLF
R_T		−0.14	−0.04	−0.30	−0.20	+0.21
K_T	−0.11		−0.07	+0.48	+0.07	−0.03
S_T	−0.12	−0.04		−0.18	−0.09	−0.06
$HCellF$	−0.29	+0.48	−0.14		+0.30	−0.30
HCF	−0.34	+0.10	−0.07	+0.36		+0.29
FLF	+0.28	−0.06	−0.04	−0.33	+0.15	

Table 6.1: Correlation coefficients between six shower shape variables calculated for signal (top right) and background (bottom left) MC samples.

The observed correlations between some variables are small enough to allow the usage of all of them in the following multivariate analysis, especially that all of them explore some separation power.

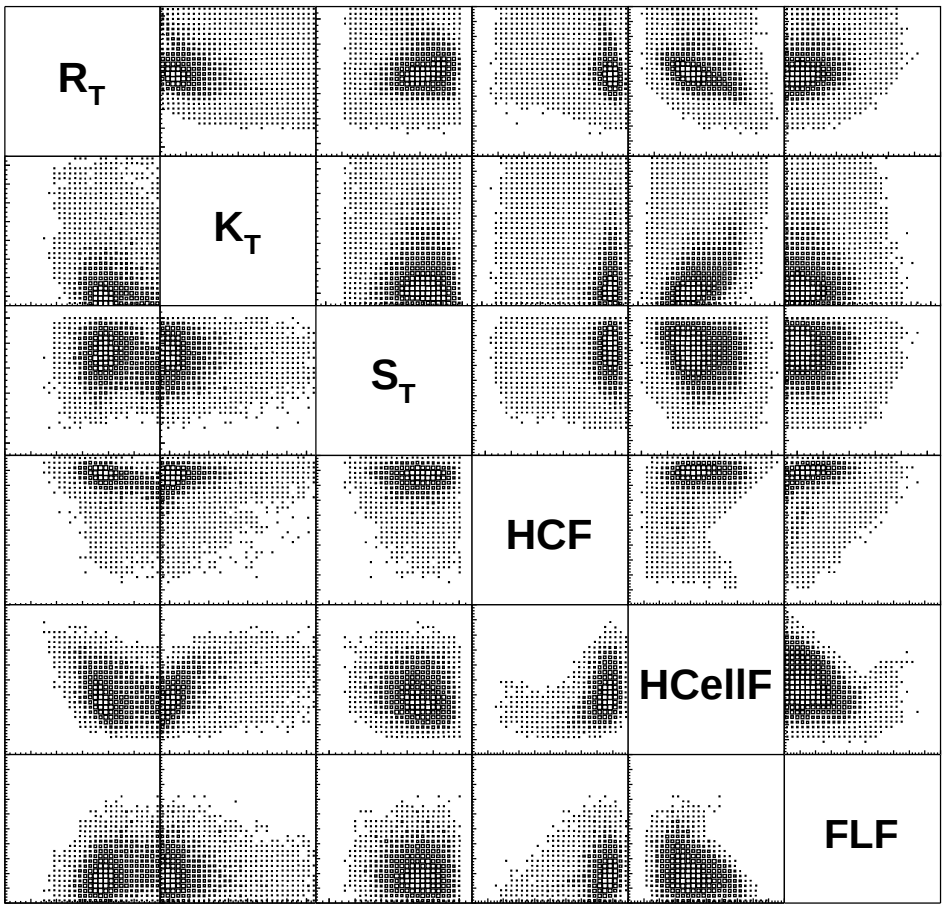

Figure 6.5: Correlation between shower shape variables for signal (top right) and background (bottom left) MC samples.

6.1.2 Separation power

To quantify the separation power hidden in each variable, one can calculate the separation $\langle S^2 \rangle$ defined for a variable x as

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{p_S(x) - p_B(x)}{p_S(x) + p_B(x)} dx \tag{6.9}$$

where $p_S(x)$ and $p_B(x)$ are probability density functions for signal and background distributions for variable x estimated by:

$$p_S(x) = \frac{dN_S(x)}{dN_S(x) + dN_B(x)}, \tag{6.10}$$

$$p_B(x) = \frac{dN_B(x)}{dN_S(x) + dN_B(x)}, \tag{6.11}$$

where dN_S and dN_B is the local number of signal and background events in a normalised x distribution.

Figure 6.6 presents the E_T dependence of the $\langle S^2 \rangle$ variable for all six shower shape variables for all considered LAr wheels. One can note the decrease of the separation power with energy for all variables but FLF. For FLF, the only input variable not based on the transverse size of the cluster, there is no physical reason to be energy dependent and in this case the separation is approximately flat in E_T. One can also see that by far, R_T variable is of the most importance, HCF and $HCellF$ being of the second importance in most of the phase space. The relatively complicated method of the symmetry calculation demands probing of the cluster shape with high resolution. The poorest granulity of **CB1** and **CB2** wheels prevents S_T variable to yield proper discrimination in those two wheels.

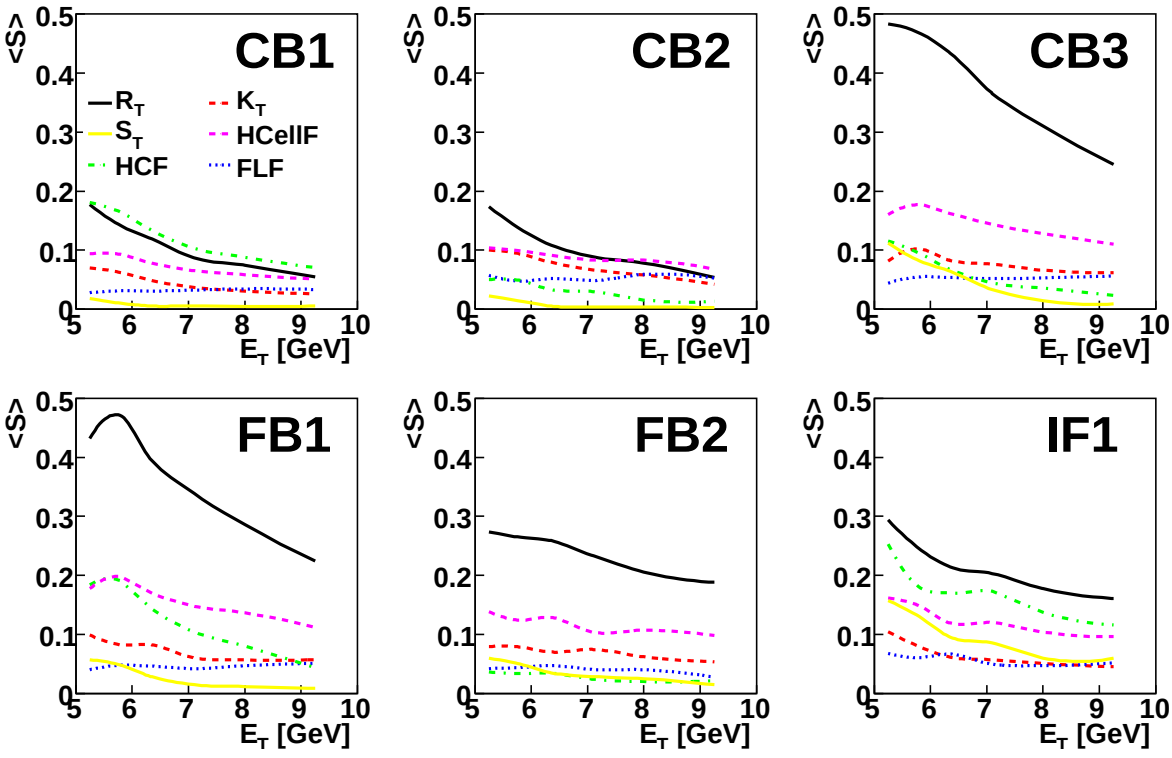

Figure 6.6: The $\langle S^2 \rangle$ dependence on E_T in six LAr wheels for R_T (solid black), K_T (dashed red), S_T (solid yellow), $HCellF$ (dashed magenta), HCF (dashed-dotted green) and FLF (dashed-dotted blue) shower shape variables.

6.2 MVA MC samples

In the section 6.1 six different shower shape variables that may be used to discriminate between signal and background were introduced. In this and following sections the method used to combine all of them into a single variable which, by using all relevant information, tries to maximise the signal to background separation power is explained. Such a variable is usually found by a classifier output of a multivariate analysis (MVA) [116].

Generally, usage of MVA consists of three steps: training, testing and evaluation. Training

step adjusts the free parameters of MVA to particular problem, the testing step checks the quality of the trained MVA and the evaluation step is used during determination of final results. Ideally, all three steps should be performed with statistically independent samples. Due to high statistical demand of the MVA, the training and testing samples in this analysis are single particle MC events with a single photons for signal (MC set XX in table 2.5) and a single hadrons for background (MC sets XXI-XXX). This simplification is possible since the MVA is purely based on the shower shapes and is insensitive to the other activity in the detector[2]. In order to minimise possible bias though, for the evaluation sample full prompt photon signal (MC sets I-II) and full background (MC set XVI) simulation is used.

Table 6.2 summarises the number of MC events used in the MVA analysis. The training sample consist of 60% randomly chosen events from the single particles MC samples. The remaining 40% define the testing sample. The full MC simulation of prompt photon and background events was used for the evaluation sample. Due to specific, highly demanding usage of the evaluated discriminator, binned in multi-dimensional, extremely fine way (see section 8.2.1), the signal evaluation sample exceed the testing sample by a factor of roughly 75. The statistics for the background evaluating sample is relatively lower, due to extremely low background selection ratio (of order 10^{-5}). Number of events listed in the discussed table differ from tables 2.3−2.5 due to the selection criteria applied for the training, testing and evaluating samples.

	training	testing	evaluating (inclusive)	evaluating (exclusive)
signal	221′709	147′456	16′981′407	11′309′443
background	246′512	164′453	72′354	58′445

Table 6.2: Number of signal and background events used for MVA.

6.3 Classifier[3]

The MVA method chosen for this analysis is commonly known as maximum likelihood or Naïve Bayes [117]. It builds a model out of probability density functions that reproduces the input variables for signal and background. For a given event i, the combined signal (background) probability $p^{MVA}_{S(B)}(i)$ is obtained by multiplying the signal (background) density probabilities $p_{S(B)}$ of all n_{var} input variables.

[2]Detailed study [66] shows that this assumption is not fully true. For a full MC event there is a certain probability for a final state soft particle being counted in the selected cluster. This effect produces statistically wider, more asymmetric and less compact clusters. In spite of this, single particle approximation has been found to be good enough for the training and testing of MVA. This approach produces a discriminator technically easier to obtain and though not anymore maximally optimal, still correct, as long as evaluation sample correctly describes real cluster shapes.

[3]All the classifiers methods were fully taken from the TMVA package [116].

$$p_{S(B)}^{MVA}(i) = \prod_{k=1}^{n_{var}} p_{S(B),k}(x_k(i)). \tag{6.12}$$

This approach assumes statistical independence of all the input variables, which is usually not fully true. The output of the MVA is a likelihood ratio, denoted further as a discriminator D. For an event i it is calculated as

$$D(i) = \frac{p_S^{MVA}(i)}{p_S^{MVA}(i) + p_B^{MVA}(i)}. \tag{6.13}$$

It can be shown that in absence of model inaccuracies (such as correlations between input variables, or an inaccurate probability density model), the ratio given by equation 6.13 provides optimal signal from background separation for the given set of input variables [].

Some other, alternative, mostly more advanced classifier definitions have been also studied for their possible application in this analysis. The result of this study is presented in appendix B together with a short discussion of their advantages and disadvantages.

The classifier has been trained independently in double differential bins of transverse energy and pseudorapidity summarised in appendix in table A-1. The choice of both dimensions has been motivated by a brief discussion in section 6.1. The discriminator distribution for both signal and background is presented in figure 6.7 as studied with the evaluating sample. The discriminator, by definition restricted to the range $0 < D < 1$, tends to peak near one for the signal events and near zero for the background events.

In addition to the separation, defined in 6.1.2, for the peaked distributions as in the case od D, one can study the significance $\langle G \rangle$ of the classifier being equal to the difference between the classifier means for signal and background divided by the quadratic sum of their root-meansquares.

$$\langle G \rangle = \int \frac{p_S(x) - p_B(x)}{p_S^2(x) + p_B^2(x)} dx, \tag{6.14}$$

The transverse energy dependence of both benchmark quantities are shown in the fig 6.8 for each LAr wheel separately. As a reference, in the same plot, the separation of the best discriminating input variable - R_T is shown. One can see the advantage of combining more than one input variable, as the discriminator separation is higher than the separation of the single input variable. There is no advantage though in the cases where the best input variable is highly dominant, as for **CB3**, **FB1** and **FB2** LAr wheels, where separation power of R_T is much higher than of all other input variables. Again, one can see the consistent drop of both, the separation and significance with the transverse energy following the dependence of the input variables.

The maximum likelihood method chosen in this analysis, is valued for its transparency, simplicity and speed. At the same time, the method's limitations are numerous. Particularly, the incorrect treatment of correlations between input variables is believed to lead to dimunision of the discrimination performance [116]. In the statistical theory there are other, more advanced classifier definitions, developed for improvement of the performance.

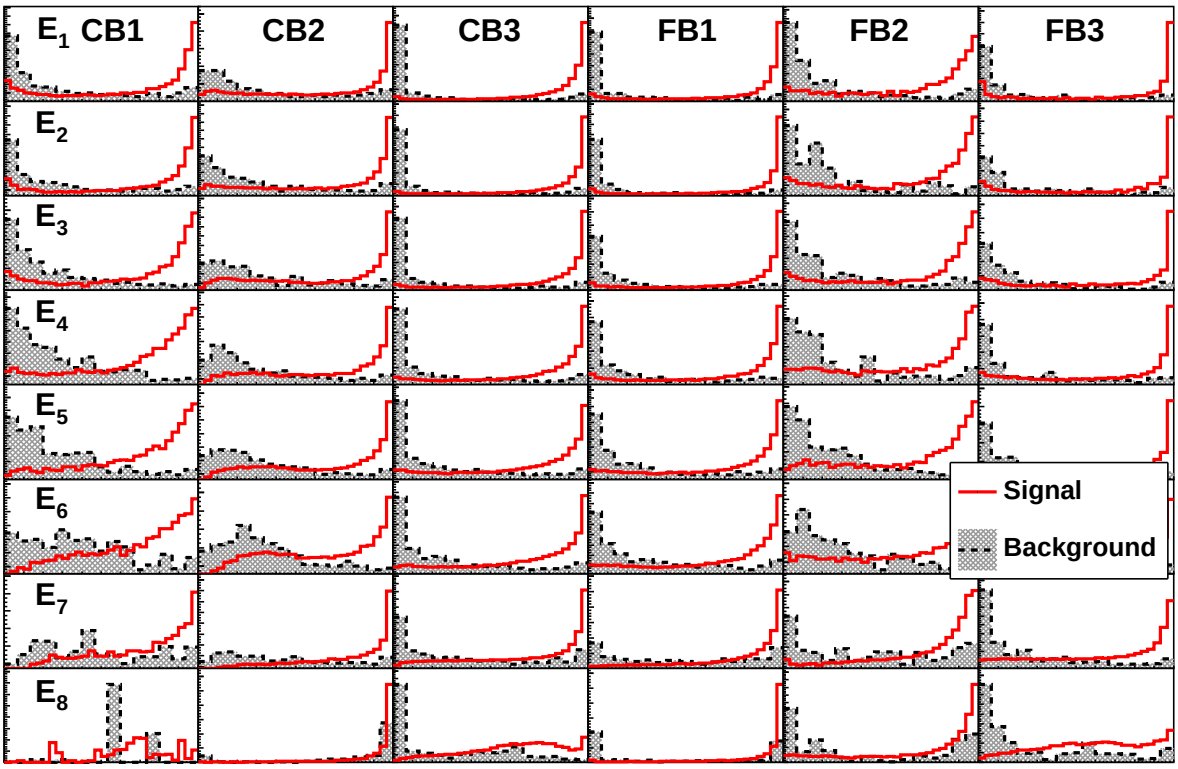

Figure 6.7: Discriminator signal and background distribution binned in a transverse energy and a pseudorapidity of the incoming particle.

The discriminator introduced in this chapter is used during the final cross section determination as an input variable for the unfolding presented in chapter 8.

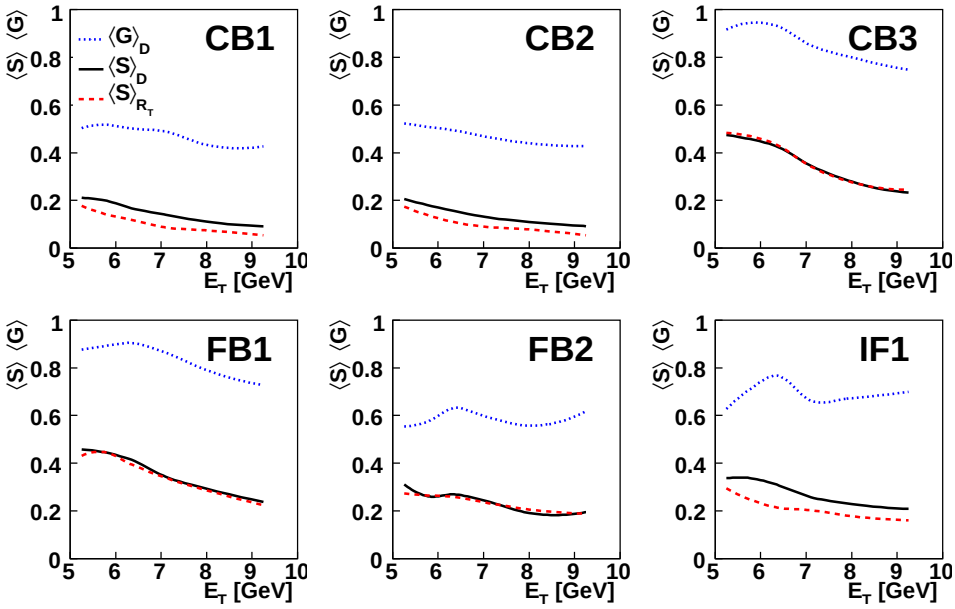

Figure 6.8: Discriminator separation (solid black) and significance (dotted blue) compared to the R_T separation (dashed red) for six LAr wheels as a function of transverse energy.

Chapter 7

Calibration and tuning

In this analysis MC is used to correct the detector level measurement to hadron level cross sections. For the correction being valid, MC events need to resemble data to high precision level and to ensure it, additional checks and calibrations explained in this chapter were performed. All the selection described in chapters 4 and 5, together with multivariate analysis introduced in chapter 6 were repeated with the corrected properties.

The most influential tuning, being the check of the cluster shape description in MC is discussed in section 7.1. Sections 7.2 and 7.3 describe electromagnetic and hadronic energy calibrations. The method of direct and resolved relative scale tuning is explained in section 7.4.

7.1 Cluster shapes tuning

The multivariable analysis introduced in chapter 6 heavily relies on the proper description of the cluster shape variables in MC. A special effort was taken to tune the MC cluster shape simulation and to correctly evaluate systematic uncertainty associated with its description.

Three data samples were used for the study of the cluster shape description:

- **BH electrons**.
  The first sample is a selection of Bethe-Heitler (BH) electons produced in a process visualised with the diagrams presented in figure 7.1a and 7.1b. Events are selected with a scattered electron in the LAr calorimeter, identified as an electromagnetic cluster with a track pointing to it, a photon in the SPACAL and nothing else in the detector.

- **DVCS photons**.
  The complementary sample to BH electrons is the selection of deeply virtual Compton scattering (DVCS) events with electron reconstructed in the SPACAL and a photon in LAr. The Feyman diagram of the DVCS event is pictured in figure 7.1c.

a. b. c.

Figure 7.1: Feynman diagrams illustrating the Bethe-Heitler events (a and b) and deeply virtual Compton scattering (c).

All the samples are statistically independent of the selection used for final results determination and provide a clean selection of electromagnetic clusters. The evaluation of the systematic uncertainty follows the method developed in [66]. MC cluster shape v is distorted by a stretching factor k:

$$v'(k) = v \cdot (1 + \frac{k}{100}), \tag{7.1}$$

and its distribution is compared to the distribution measured in the data. An example of the comparison is presented in figure 7.2, where the FLF MC distribution in wheel **CB1** is stretched against the data distribution measured in the BG sample. One may see that for negative k the MC distribution is shifted towards too low FLF values, while for positive k in the opposite direction. For k close to zero, both distributions are comparable, being at it best for $k \approx 2$.

For each factor k the χ^2 can be calculated between the data histogram and the stretched MC distribution:

$$\chi^2 = \sum_i (b_i^D - b_i^{MC})^2/(\sigma_{D,i}^2 + \sigma_{MC,i}^2), \tag{7.2}$$

where b_i^D and b_i^{MC} denote the data and MC histogram content for bin i, σ its error and the sum runs over all non-empty histogram bins. The number of non-empty bins correspond to the number of degrees of freedom (NDF).

Figure 7.2 shows the $\chi^2(k)/NDF$ dependence for all six shower shape variables in **CB1** wheel as studied with the BG sample. A clear minima close to $k = 0$ can be observed in each case. The $\min(\chi^2)$ is used for the correction applied to the MC simulation of the cluster shapes, while $\min(\chi^2) \pm NDF$ for the evaluation of its uncertainty. In case of DVCS and BH samples, error deduced by the χ^2 method is dominated by statistical error of both samples, which leads to an overestimated uncertainty of the cluster shapes description.

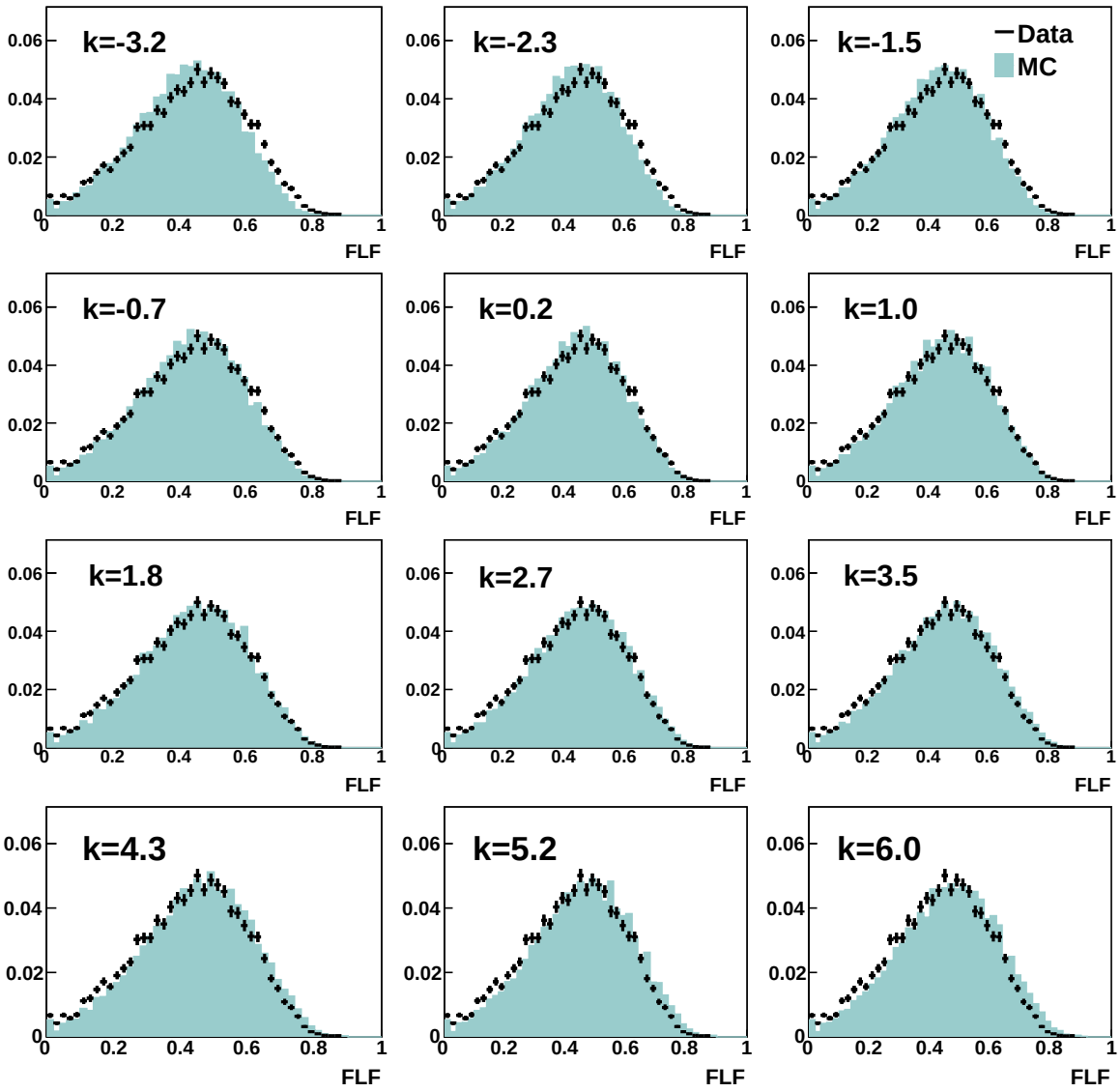

Figure 7.2: The FLF variable description in MC studied in **CB1** wheel with BG sample. The stretching factor k applied to the MC sample is printed in each plot.

The final estimate of the k factors is done with the BG sample by χ^2 method additionally verified by careful examination of the compared histograms. Since the background sample is produced by superimposition of two or more single photon showers, the cluster shape misrepresentation in this case is expected to be only higher than in a single photon sample. Accordingly, it was decided that such determined uncertainties may be used for signal selection. The comparison of all three samples in regions with high enough statistics lead to comparable corrections with background sample having the clear statistical advantage.

The derived stretching factors k, used for correction applied to MC cluster simulation, with their uncertainties are summarized in table 7.1.

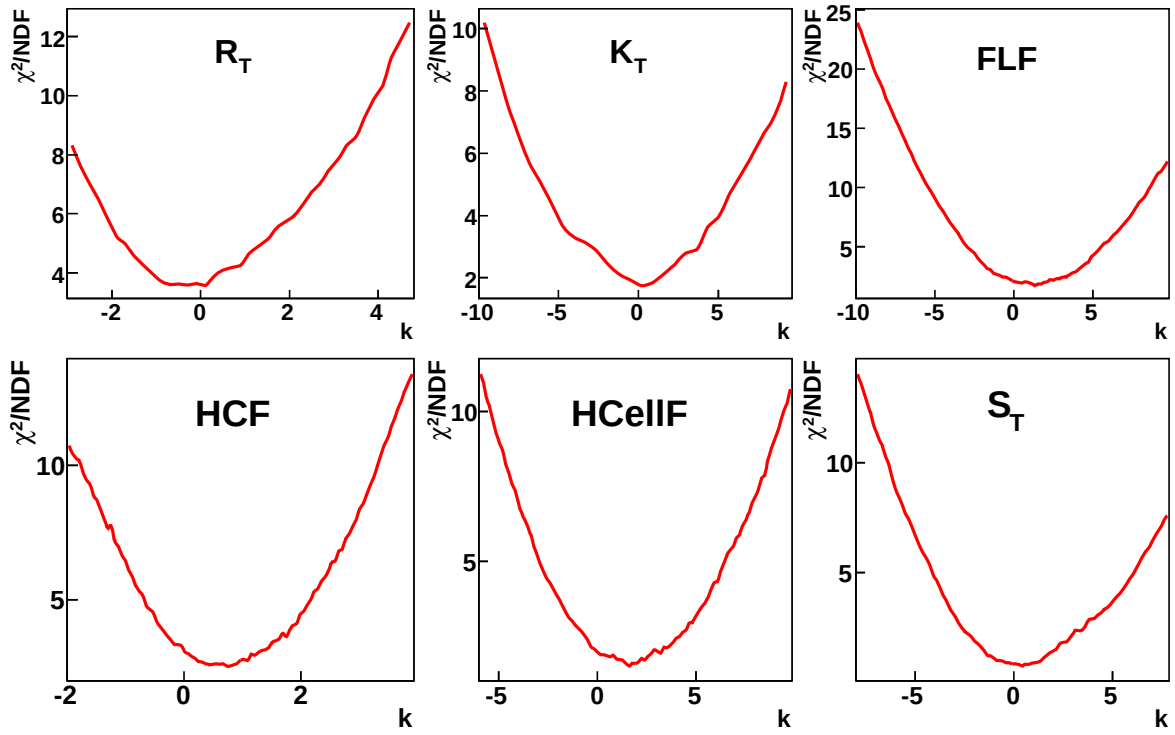

Figure 7.3: The χ^2/NDF dependence on the stretching factor k for six cluster shape variables in the **CB1** wheel for the background enhanced sample.

	CB1	**CB2**	**CB3**	**FB1**	**FB2**	**IF1**
R_T	-0.5 ± 1.5	0.0 ± 1.0	-1.0 ± 1.5	-1.5 ± 1.5	0.0 ± 4.0	$+2.0 \pm 8.0$
K_T	$0.0 \pm +4.0$	-1.0 ± 3.0	-0.5 ± 2.5	$+1.0 \pm 2.0$	0.0 ± 5.0	$+6.0 \pm 10.0$
S_T	$+0.2 \pm 1.8$	0.0 ± 1.5	$+1.0 \pm 2.0$	$+1.5 \pm 2.5$	$+2.0 \pm 4.0$	0.0 ± 5.0
$HCellF$	$+1.0 \pm 1.5$	0.0 ± 1.5	0.0 ± 1.5	-0.8 ± 2.2	-5.0 ± 5.0	$+5.0 \pm 7.5$
HCF	$+0.5 \pm 1.0$	$+0.5 \pm 0.5$	$+0.2 \pm 0.5$	$+0.5 \pm 1.0$	$+2.0 \pm 4.0$	-1.5 ± 3.5
FLF	$+1.0 \pm 2.0$	$+0.5 \pm 2.0$	$+1.5 \pm 2.5$	$+5.0 \pm 5.0$	$+10.0 \pm 10.0$	$+4.0 \pm 10.0$

Table 7.1: Stretching factors k and they uncertainties as determined for six cluster shapes.

7.2 Photon energy calibration

The energy calibration of the LAr calorimeter was originally calibrated using test beam measurements [102, 103]. Since then, more precise calibration using real data was performed [118] with electrons in high Q^2 neutral current events. This provides a good calibration at higher energies.

Since analysis uses low energetic clusters, the energy calibration factors were cross checked

using already introduced electrons from Bethe-Heitler events and photons from deeply virtual Compton scattering events, in a similar way to [66]. For BH electrons, the transverse LAr energy was calibrated with the help of the electron transverse momentum measured directly in the tracker. For DVCS photons the energy can be calculated indirectly using the double angle method:

$$E^{\gamma}_{T,DA} = \frac{2E_e \sin\theta^{had} \sin\theta^{e}}{\sin\theta^{had} + \sin\theta^{e} - \sin(\theta^{had} + \theta^{e})}, \tag{7.3}$$

where E_e is the electron beam energy, θ^e is the polar angle of the scattered electron and θ^{had} is the inclusive hadronic angle. Both methods confirm the validity of the calibration factors used in the analysis, within statistical significance.

A similar study was performed for the calibration of photons in the MC simulation. Figure 7.4 shows the comparison between MC generated E_T^{gen} and reconstructed E_T^{rec} transverse energy additionally binned in the pseudorapidity of the photon. The mean value of the fitted gaussian resolution function corresponds to typically two percent miscalibration.

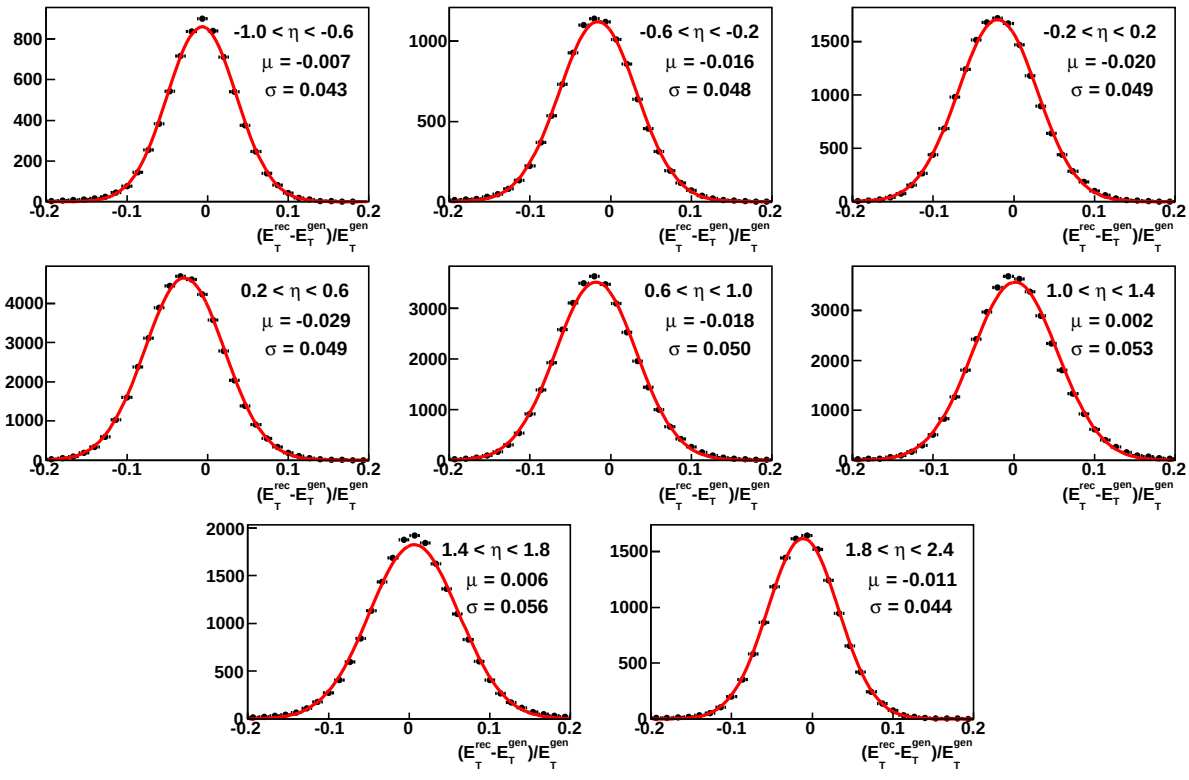

Figure 7.4: Transverse energy calibration studied with the help of reconstructed energy E_T^{rec} and generated energy E_T^{gen} in eight bins of pseudorapidity. Distributions are consistent with a Gaussian function with fitted mean value μ and standard deviation σ. Both parameters are printed in the respective subfigures.

Comparison to the data miscalibration, together with other H1 analyses (particularly DVCS study of [119]) lead to a relative MC misrepresentation of the data electromagnetic

energy on the level of 1% in central LAr wheels and 4% in the most forward LAr wheel. No corrections are applied to the MC simulation and values above are used for evaluation of the systematic error.

7.3 Hadronic energy calibration

An additional offline energy calibration correction known as *HighPtJetCalibration* [120] has been applied to the hadronic energy reconstructed in both MC and data. Figure 7.5 shows the mentioned distribution for an exclusive event selection for jets before (a) and after (b) *HighPtJetCalibration*. The significant improvement in MC description of jet transverse energy in the exclusive sample (see section 5.3) can be observed.

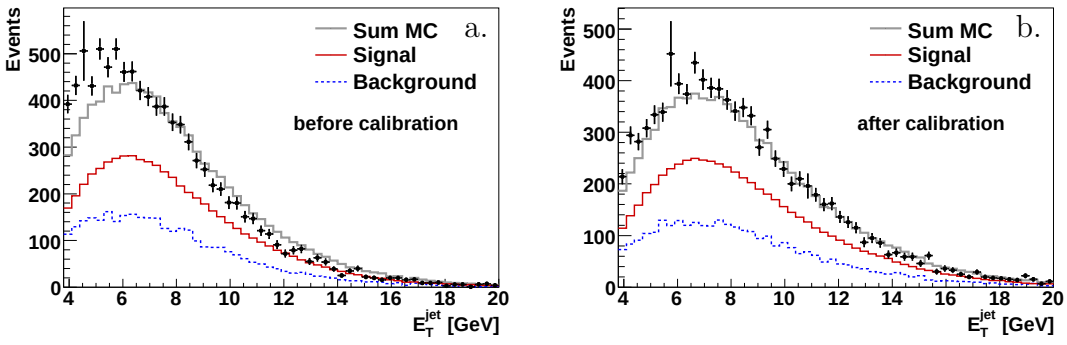

Figure 7.5: Transverse energy of the hadronic jet in the exclusive selection before (a) and after (b) *HighPtJetCalibration* calibration.

The calibration has been verified using the same exclusive prompt photon selection. With the assumption of small intrinsic parton k_T one could expect prompt photon being fully balanced in transverse energy with the sum of the remaining hadronic final state. Thus the ratio between the transverse energy of the hadronic final state E_T^{HFS} and the transverse energy of the photon candidate E_T^{γ} has been studied and presented in figure 7.6 as a function of E_T^{γ} and η^{γ}. One can see significant improvement in both MC and data after the calibration.

The double ratio plots are commonly used to check the quality of the MC simulation. The double ratio variable D_{Ej} is calculated as

$$D_{Ej} = \left(\frac{E_T^{HFS}}{E_T^{\gamma}}\right)_{Data} / \left(\frac{E_T^{HFS}}{E_T^{\gamma}}\right)_{MC} \tag{7.4}$$

and measures the hadronic energy model inaccuracy. Double ratio plots in bins of E_T^{γ} and η^{γ} are plotted in figure 7.7. One can conclude that MC describes jet energy within 2% uncertainty and this value, also consistent with other H1 studies had been taken for the evaluation of systematic error.

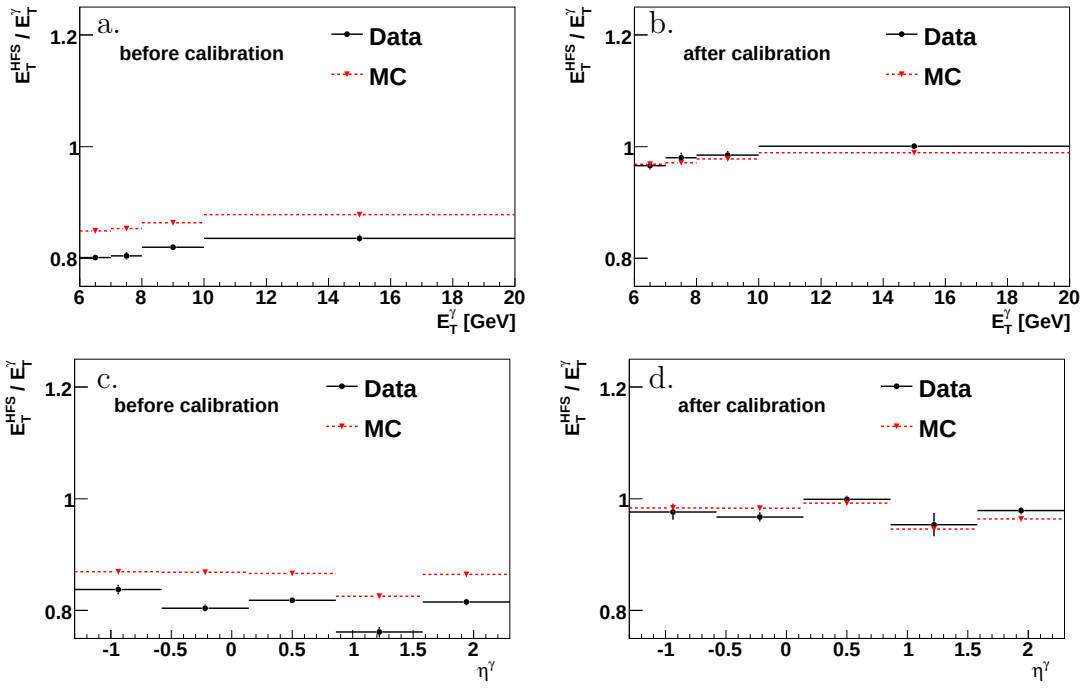

Figure 7.6: The ratio of transverse energy of the hadronic final state to transverse energy of prompt photon candidate as a function of transverse energy of photon candidate E_T^{γ} (a, b) and its pseudorapidity η^{γ}(c, d). Plots show situation before (a, c) and after (b, d) hadronic energy calibration.

7.4 Event class ratios

The ratio of direct to resolved cross section in MC is tuned to the measurement. In this section, a double-step method of tuning that ratio to data is explained. The ratio of direct to resolved events scale is important due to difference in the selection efficiency between the two samples, being on average equal to 0.34 for a direct prompt photon signal and 0.27 for a resolved prompt photon signal. Since the efficiency correction is taken purely from MC, a proper sample mix is required. For the purpose of the event class ratio determination, the distributions of x_{γ}^{jet} observable is used. Its definition being similar to the x_{γ}^{LO} given by equation 1.41 is an another estimator of the x_{γ} variable, but since it is using two leading jets, is valid as well for background events:

$$x_{\gamma}^{jet} = \frac{E_T^{jet1}\exp(-\eta^{jet1}) + E_T^{jet2}\exp(-\eta^{jet2})}{2yE_e}, \tag{7.5}$$

with E_T^{jet} and η^{jet} being the transverse energy and pseudorapidity of the two leading jets, E_e is the energy of the electron beam and y is the inelasticity estimator. For prompt photon events one of the jets is the photon jet.

- **Tuning background event class ratios**
  The background scales are studied using the low isolated photons (with the isolation

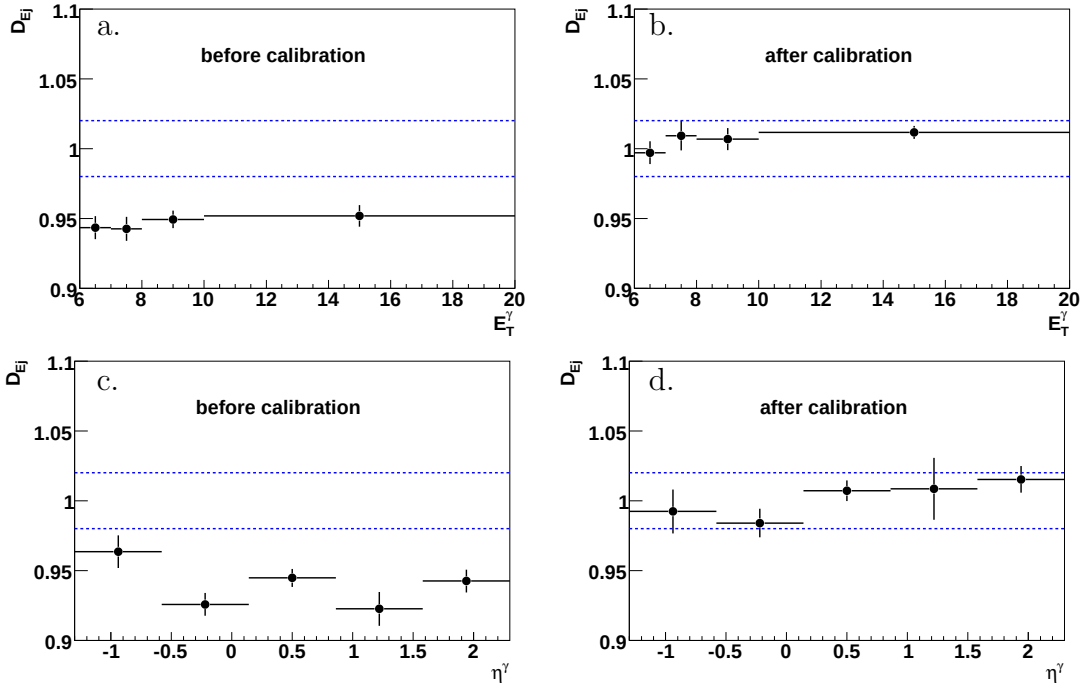

Figure 7.7: Double ratio plots as a function of transverse energy of photon candidate E_T^γ (a, b) and its pseudorapidity η^γ (c, d). Plots show situation before the calibration (a, c) and after calibration (b, d). Horizontal lines show the evaluated 2% error on the MC description of hadronic energy.

criteria $0.7 < z < 0.9$). In such a phase space, the prompt photon signal contributes on the negligible level and the background samples may be scaled to the measured data distribution. Figure 7.8a presents the x_γ^{jet} distribution together with the fitted direct and resolved background MCs. Significant difference between both distributions can be observed, which allows a meaningful fit. A good description of the data by the sum of scaled background MC can be noted as well. The background direct MC must be scaled by a factor 5.46, and the resolved component by 1.71. One should though keep in mind that due to preselection of the isolated neutral particle on the generator level (see section 2.1.2) those scaling factors do not have much meaning outside of the analysis presented in this particular thesis. The background scaling factors, with the assumption of the cluster shapes insensitivity to event class, is important only for the next step of the event class ratio tuning.

- **Tuning signal event class ratios**
  The signal event class ratio determination is more complicated, as it is not easy to obtain pure enough signal event sample. The signal enhanced selection follows the standard inclusive prompt photon selection with an additional cut on the discriminator ($D > 0.7$). In this way the influence of the background is minimised, but still can not be fully neglected. The x_γ^{jet} distribution together with two fitted components (direct and resolved prompt photons) and a fixed background distribution is shown in figure 7.8b. The background in this case was fixed to the scales obtained

in the previous step. One may see that in this particular selection phase space, background and resolved signal components have a similar distribution and can not be simultaneously fitted. The fitted scaling factor for a direct prompt photon sample is equal to 1.18 ± 0.03, while the resolved component needs a higher scaling of 1.57 ± 0.08. One should note that what influences the final measurement is only the relative direct to resolved scale. The uncertainty of the fitted scales were used for the evaluation of the systematic error associated with the effect explained above.

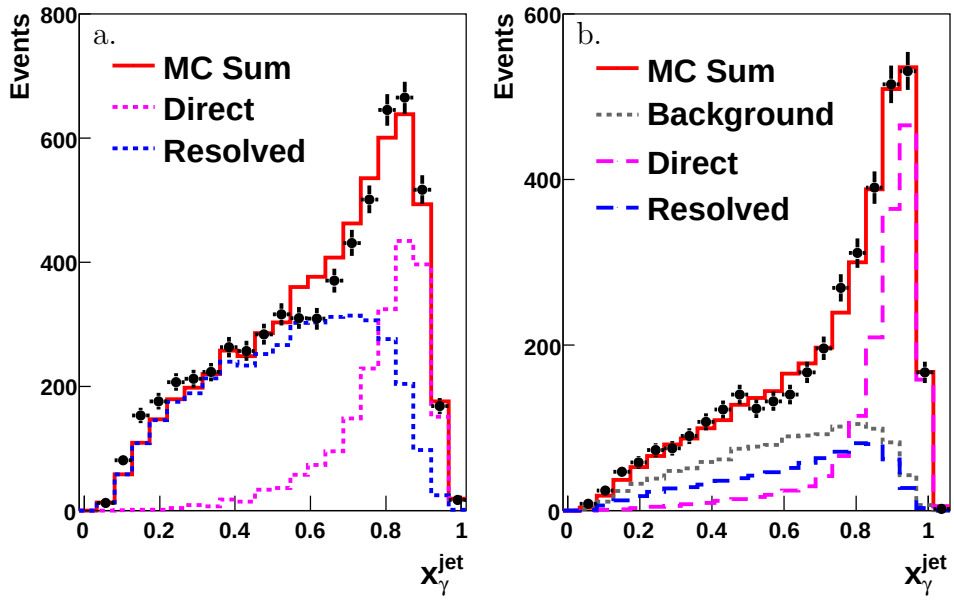

Figure 7.8: Event class ratios in the background enhanced phase space (a) and signal enhanced phase space (b).

Chapter 8

Cross section building

This chapter gives the detailed information about the method used to build the cross sections with a special emphasis on the unfolding procedure. The unfolding problem and solution is theoretically formulated in section 8.1 while the application is given in section 8.2. Evaluation of an error matrix and the treatment of systematics is summarised in section 8.3. The final cross sections building using the unfolding output is explained in section 8.4. The validation of the cross section extraction procedure by the means of two major consistency checks is presented in sections 8.5 and 8.6. Finally, section 8.7 describes the method used to build comparable cross sections out of the QCD calculation output.

8.1 Unfolding problem

In the following section the unfolding problem is formulated in the most general way. Commonly known issues are addressed and possible solutions are given.

In high energy experiments the common problem is the extraction of a signal distribution $f^{true}(x_{gen})$ of a physical quantity x_{gen}. Due to various detector effects the true $f^{true}(x_{gen})$ distribution differs from the measured $g^{obs}(y_{det})$ distribution of the corresponding directly observed variable y_{det}. Mathematically, both true and measured distributions are related by the equation

$$g^{obs}(y_{det}) = \int A(x_{gen}, y_{det}) f^{true}(x_{gen}) dx_{gen}, \tag{8.1}$$

with $A(x_{gen}, y_{det})$ being the detector response function that incorporates effects such as detector resolution, acceptance and efficiency.

Practically, in most of the cases g^{obs} and f^{true} distributions are represented by histograms (vectors), and the continuous $A(x_{gen}, y_{det})$ becomes a migration matrix $\mathbf{A}$. In such a case the equation 8.1 can be represented by its matrix form:

$$\vec{y}_{obs} = \mathbf{A}\vec{x}_{true}. \tag{8.2}$$

The correctly normalised migration matrix has an easy interpretation, its element $\mathbf{A}_{ij}$ represents the probability of an event originating from bin j of $\vec{x}_{true}$ to be measured in bin i of $\vec{y}_{obs}$. The sizes of $\vec{y}_{obs}$ and $\vec{x}_{true}$ vectors do not need to be equal, and y_{obs} does not need to be an estimator of x_{true}. In most applications though, it is the case, since $\vec{y}_{obs}$ and $\vec{x}_{true}$ should be correlated if a reasonable solution is expected. A large size of $\vec{y}_{obs}$ usually leads to improvement of the unfolding performance [121]. For that reason, it is requested that for a correct unfolding setup $dim(\vec{y}_{obs}) \geq 2 \times dim(\vec{x}_{true})$. The migration matrix is determined using MC simulation that incorporates most of the detector effects.

The unknown $\vec{x}_{true}$ distribution is usually found by minimising the $\chi^2_{\mathbf{A}}$ function with respect to $\vec{x}_{true}$, where $\chi^2_{\mathbf{A}}$ measures the difference between both sides of equation 8.2:

$$\chi^2_{\mathbf{A}} = 1/2 \cdot (\vec{y}_{obs} - \mathbf{A}\vec{x}_{true})^T \mathbf{V}^{-1} (\vec{y}_{obs} - \mathbf{A}\vec{x}_{true}). \tag{8.3}$$

The matrix $\mathbf{V}^{dim(\vec{y}_{obs}) \times dim(\vec{y}_{obs})}$, being the $\vec{y}_{obs}$ covariance matrix, is estimated using the measured errors of the $\vec{y}_{obs}$.

8.1.1 Regularisation

It has been observed [121] that the unfolding in its purest form leads to high $\vec{x}_{true}$ fluctuations with high anti-correlations between neighboring bins. The problem can be partially removed by bin averaging procedure. Another possibility is to use output regularisation incorporated into the unfolding procedure that implies some sort a'priori chosen $\vec{x}_{true}$ form. Any deviation from this form introduce a penalty to the minimised χ^2 function.

In the most general case, the arbitrarily defined $\vec{x}_{true}$ regularisation condition $\mathbf{L}$ introduces the $\chi^2_{\mathbf{L}}$ penalty

$$\chi^2_{\mathbf{L}} = \vec{x}^T_{true}\ \mathbf{L}\ \vec{x}_{true} \tag{8.4}$$

and may enter the minimised χ^2 function in the following way:

$$\chi^2 = \chi^2_{\mathbf{A}} + \tau^2 \cdot \chi^2_{\mathbf{L}}, \tag{8.5}$$

where τ is the regularising parameter adjusting the scale of the regularisation penalty.

The regularisation condition $\mathbf{L}$ can take a variety of forms, with three most common ones being the following [122]:

- **Regularisation on the size**
  The easiest type of the regularisation introduces a penalty proportional to the square of the size of the unfolded $\vec{x}_{true}$ distribution. The regularisation tries to enforce low and peaks deprived $\vec{x}_{true}$ distribution. The regularisation condition $\mathbf{L}$ adopts than its simplest form:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{8.6}$$

- **Regularisation on the derivative**
  The more advanced regularisation tries to enforce a flat $\vec{x}_{true}$ distribution, without favouring low solutions. It is achieved by minimising the derivative of the $\vec{x}_{true}$ distribution. The first derivative for point i of the $\vec{x}_{true}$ distribution may be approximated by

$$x'^{\,i}_{true} \approx (x^{i+1}_{true} - x^{i}_{true})/\Delta x^i \tag{8.7}$$

  where Δx^i is the width of the bin i. In case of equidistant bin widths equal to one, the regularisation condition $\mathbf{L}$ takes the form:

$$\mathbf{L} = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 \end{bmatrix}. \tag{8.8}$$

- **Regularisation on the curvature**
  Even more advanced regularisation is trying to enforce a smooth monotonic shape of $\vec{x}_{true}$ by minimising its second derivative. The second derivative for point i can be calculated similarly to the previous case using the Taylor series expansion as:

$$x''^{\,i}_{true} \approx (x^{i-1}_{true} - 2x^{i}_{true} + x^{i+1}_{true})/(\Delta x^i)^2. \tag{8.9}$$

  With all the bin widths equal to one $\mathbf{L}$ adopts form:

$$\mathbf{L} = \begin{bmatrix} -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -2 \end{bmatrix}. \tag{8.10}$$

Minimisation of the χ^2 function (equation 8.5) can be performed analytically by solving the matrix equation $\partial\chi^2/\partial\vec{x}_{true} = 0$. The minimum of χ^2 is achieved for

$$\vec{x}_{true} = \mathbf{E_u}\mathbf{A}^T\mathbf{V}\vec{y}_{obs} \tag{8.11}$$

with

$$\mathbf{E_u} = (\mathbf{A}^T\mathbf{V}\mathbf{A} + \mathbf{L}^2\tau^2)^{-1} \tag{8.12}$$

where $\mathbf{E_u}$ is the $\vec{x}_{true}$ covariance matrix and is used to estimate the final error on the determined solution.

The regularising parameter τ is a free parameter which needs to be properly adjusted before the final calculation can be performed. Two τ determination methods, the so called L-curve method and minimal correlation method are presented below. Both involve solving equation 8.11 for the whole range of regularisation parameters τ and choosing the most appropriate one.

- **L-curve method** [123]
  Calculated solutions for the whole range of regularisation parameters τ can be represented by points on the $(\chi^2_{\mathbf{A}}, \chi^2_{\mathbf{L}}/\tau^2)$ plane. Solutions are expected to allocate along the curve plotted in figure 8.1. For small values of τ, equation 8.5 is fully dominated by $\chi^2_{\mathbf{A}}$ and solutions can be found in the top left corner of the plot. Those solutions are characterised by low $\chi^2_{\mathbf{A}}$ values (that is the real χ^2 in the strict statistical usage of the term), but are not regularised and might be affected by strong fluctuations. With rising regularisation, the solutions migrate to lower parts of the plot due to the regularisation parameter τ in the denominator of the vertical axis definition. The solutions give constant $\chi^2_{\mathbf{A}}$ until the regularisation part $\chi^2_{\mathbf{L}}$ of the equation 8.5 becomes comparative and at some point even dominant. With still rising τ the solution is fully driven by the $\chi^2_{\mathbf{L}}$ producing unreasonable $\chi^2_{\mathbf{A}}$. The optimal choice of the regularisation is the point with still low $\chi^2_{\mathbf{A}}$, but already regularised. It is indicated by the arrow in the figure 8.1.

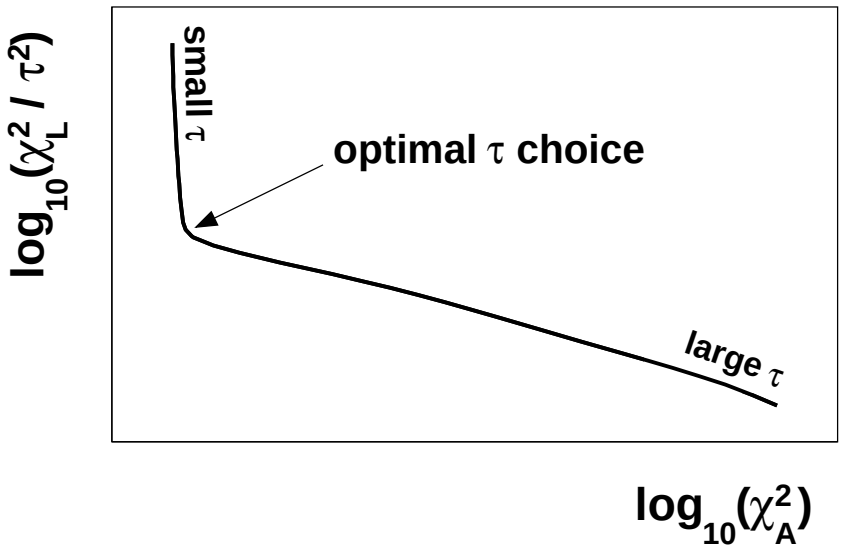

Figure 8.1: The expected L-curve shape used for the determination of the regularisation parameter τ. The optimal choice of τ is indicated by an arrow.

- **Minimal correlation method**
  The minimal correlation method makes use of interesting features of the derivative and the curvature regularisation schemes. Both of those methods connect neighboring $\vec{x}_{true}$ bins and therefore have a tendency to create a positive correlation. For those regularisation schemes, the regularisation parameter τ can be tuned such as to introduce the smallest possible correlations.

If applicable, both methods choose comparative regularisation parameter τ.

8.2 Unfolding application

Having the short theoretical introduction to the unfolding problem, this section explains the way the unfolding is treated in this particular analysis.

The cross sections binned in the interesting variables (see section 1.5) are obtained by bin averaging (see section 8.4.1) of higher dimensional unfolding output, so more than one cross section can be actually derived from a single unfolding. Table 8.1 associates all the final cross sections with their codes[1] Every *unfolding code*, defined in the same table, corresponds to one independent unfolding solution. All the variables and phase spaces are defined in section 1.5.

Phase space	Variable	Cross Section Code	Unfolding Code
inclusive	E_T^γ	**i**	**A**
inclusive	η^γ	**ii**	**A**
exclusive	E_T^γ	**iii**	**B**
exclusive	η^γ	**iv**	**B**
exclusive	E_T^{jet}	**v**	**C**
exclusive	η^{jet}	**vi**	**D**
exclusive	x_γ^{LO}	**vii**	**E**
exclusive	x_p^{LO}	**viii**	**F**
direct	$p_\perp$	**ix**	**G**
direct	$\Delta\phi$	**x**	**H**
resolved	$p_\perp$	**xi**	**I**
resolved	$\Delta\phi$	**xii**	**J**

Table 8.1: The definition of cross sections codes and unfolding codes.

8.2.1 Migration matrix

In this analysis, unfolding plays an essential role. It addresses features of the analysis such as correct treatment of bin to bin migration effects, detector acceptance and efficiency corrections as well as signal and background discrimination. The migration matrix, being the most important unfolding component, has been carefully developed.

The main advantage of the unfolding procedure is its relatively high reliability in the presence of existing model inaccuracies, particularly in case of high bin to bin migrations. The model independency can be achieved only for the variables used as unfolding output, so it is important to correctly select relevant variables and even, if possible, use multi dimensional binning. Table 8.2 lists all the unfolding codes with the chosen input

[1] For a simplification, through out of this chapter cross sections codes are used.

(**Reconstructed bins**) and output (**Signal bins**) coded binning. The actual bin edges are given in appendix A (tables A-2 − A-5).

Unfolding Code	Reconstructed bins	Signal bins
A, B	$10(E^{\gamma}_{T,IN10}) \times 10(\eta^{\gamma}_{IN10}) \times 5(D_{IN})$	$4(E^{\gamma}_{T,OUT4}) \times 5(\eta^{\gamma}_{OUT5})$
C	$8(E^{jet}_{T,IN}) \times 5(E^{\gamma}_{T,IN5}) \times 8(\eta^{\gamma}_{IN8}) \times 5(D_{IN})$	$4(E^{jet}_{T,OUT}) \times 2(E^{\gamma}_{T,OUT2}) \times 5(\eta^{\gamma}_{OUT5})$
D	$8(\eta^{jet}_{IN}) \times 5(E^{\gamma}_{T,IN5}) \times 8(\eta^{\gamma}_{IN8}) \times 5(D_{IN})$	$4(\eta^{jet}_{OUT}) \times 2(E^{\gamma}_{T,OUT2}) \times 5(\eta^{\gamma}_{OUT5})$
E	$10(x_{\gamma,IN}) \times 5(E^{\gamma}_{T,IN5}) \times 8(\eta^{\gamma}_{IN8}) \times 5(D_{IN})$	$5(x_{\gamma,OUT}) \times 2(E^{\gamma}_{T,OUT2}) \times 5(\eta^{\gamma}_{OUT5})$
F	$8(x_{p,IN}) \times 5(E^{\gamma}_{T,IN5}) \times 8(\eta^{\gamma}_{IN8}) \times 5(D_{IN})$	$4(x^{p,OUT}) \times 2(E^{\gamma}_{T,OUT2}) \times 5(\eta^{\gamma}_{OUT5})$
G, I	$9(p_{\perp,IN}) \times 5(E^{\gamma}_{T,IN5}) \times 8(\eta^{\gamma}_{IN8}) \times 5(D_{IN})$	$4(p_{\perp,OUT}) \times 2(E^{\gamma}_{T,OUT2}) \times 5(\eta^{\gamma}_{OUT5})$
H, J	$7(\Delta\phi_{IN}) \times 5(E^{\gamma}_{T,IN5}) \times 8(\eta^{\gamma}_{IN8}) \times 5(D_{IN})$	$4(\Delta\phi_{OUT}) \times 2(E^{\gamma}_{T,OUT2}) \times 5(\eta^{\gamma}_{OUT5})$

Table 8.2: The definition of input and output binning used for migration matrix definition. Here only numbers of bins and binning codes are given, the actual bin edges are listed in appendix A.

Each of the input bins is additionally subbinned in the discriminator (D_{IN}) defined in section 6.3. That allows the determination of the signal content in every of the input bins during unfolding. As it was already noted, the discriminator, built on shower shape variables, depends on E^{γ}_{T} and η^{γ} of the studied photon candidate. For that reason, each input binning, even when not directly needed, is binned in both observables with ten η^{γ} bins (eight for more complex unfoldings) sensitive to the LAr calorimeter wheel edges and ten (five) E^{γ}_{T} bins. A similar, though less fine binning is chosen for the output binning. In this way, not only the discriminator is properly taken into account, but also unfolding becomes relatively insensitive to the possible model bias introduced along the η^{γ} and E^{γ}_{T} directions. Every output bin set, if possible, includes the underflow and overflow bin to correctly treat the migration from outside of the hadron level phase space.

The migration matrix is build with the help of signal MC (MC sets I-II in table 2.3). The **Reconstructed bin** is chosen for events passing the selection criteria (summarized in table 5.1), while **Signal bin** is defined for events from within the measurement phase space (tables 1.4 and 1.5) on the hadron level. The main part of the migration matrix body is filled with events both generated in the phase space and passing selection criteria. The treatment of other events is discussed later in this chapter.

The high goals set for the unfolding procedure already in the beginning of this section are reflected in the relative complexity of the migration matrix. The simplified sketch of the matrix can be seen in figure 8.2. The **Input** side of the matrix is drawn horizontally, where in addition to the **Reconstructed bins** (here in this simplified picture with $5D \times 3E^{\gamma}_{T} \times 2\eta^{\gamma}$ bins) so called **Side bins** are included. Similarly, along the vertical **Output** direction

Migration bins and **Background bins** add the complexity to already explained **Signal bins**.

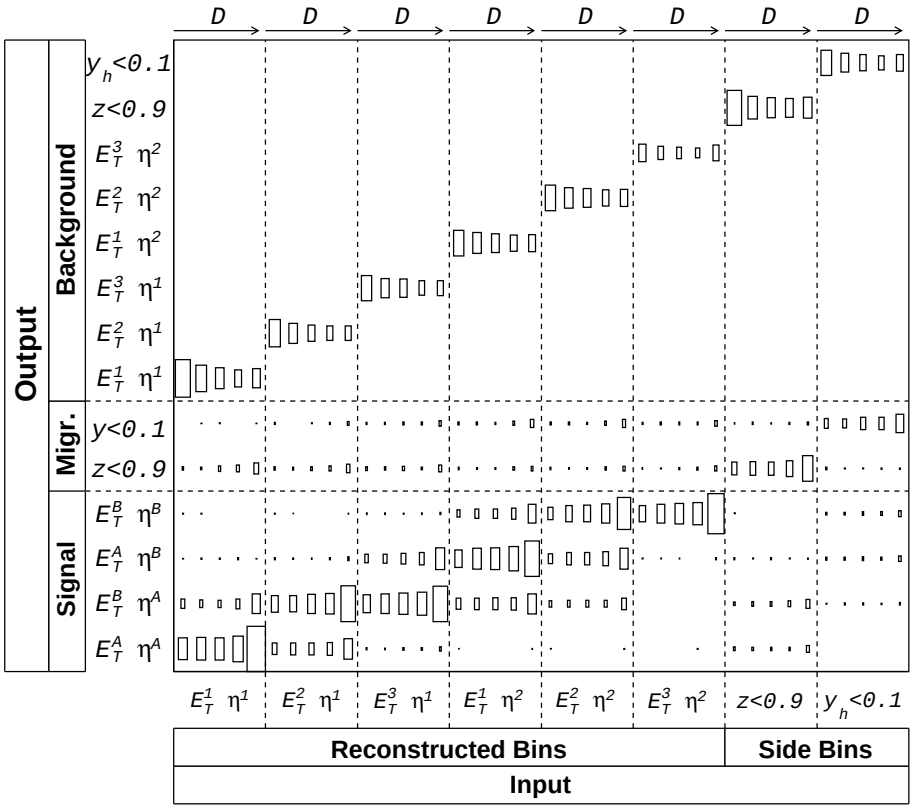

Figure 8.2: The schematic representation of the migration matrix used in this analysis. Unfolding **Input** is build of arranged **Reconstructed Bins** and **Side Bins**, while the **Output** consists of **Signal**, **Migration** and **Background Bins**. The meaning of all the bin sets is explained in the text.

The unfolding matrix is the product of two paradigms:

- Every event that fulfils the hadron level phase space criteria (tables 1.4 and 1.5) or detector level phase space criteria (table 5.1), needs to be accommodated in the unfolding matrix exactly once.

- Every **Output** bin needs to be linearly independent of all the other **Output** bins in order to avoid ambiguity in the unfolding problem[2].

[2]The ambiguity is usually signalled by the singularity of matrix $(\mathbf{A}^T\mathbf{V}\mathbf{A} + \mathbf{L}^2\tau^2)$ and inability to calculate the $\mathbf{E_u}$ (see equation 8.12)

8.2.1.1 Migration bins

The first paradigm implies the incorporation of the **Migration bins** into the output direction of the unfolding to accommodate events generated outside of the hadron level phase space, but migrating into the detector level selection. Those events are expected to be measured in one of the data bins and need to be subtracted. Different hadron level cuts introduce their own migrations into the selection and are handled separately. This allows subtraction even if the MC does not fully describe relevant variables. Each migration bin is additionally subdivided into $2E_T^\gamma \times 3\eta^\gamma$ bins[3] (listed in appendix table A-6) to minimise the possible model bias.

Table 8.3 summarises the choice of the migration bins used for different unfoldings. The inclusive measurement uses only migrations due to the isolation cut and inelasticity cut on hadron level. The exclusive measurements includes in addition the migrations caused by cuts on jet properties. The migration due to the cut on transverse energy of the jet in case of unfolding **C** is already taken into account in the relevant underflow **Signal bin**, so the corresponding **Migration bin** can be skipped. A similar situation occurs for unfolding **D** and migrations due to cuts on pseudorapidities of the hadronic jet. For the same reason there are no special migration bins handling the cuts on transverse energy and pseudorapidity of the photon since those underflow / overflow bins are always present among the **Signal bins**.

Migration bin	$z < 0.9$	$y < 0.1$	$y > 0.7$	$E_T^{jet} < 4.5\text{GeV}$	$\eta^{jet} < -1.3$	$\eta^{jet} > 2.3$	$x_\gamma^{LO} < 0.8$	$x_\gamma^{LO} > 0.8$
A	+	+	+	-	-	-	-	-
B, E, F	+	+	+	+	+	+	-	-
C	+	+	+	-	+	+	-	-
D	+	+	+	+	-	-	-	-
G, H	+	+	+	+	+	+	+	-
I, J	+	+	+	+	+	+	-	+

Table 8.3: The **Migration bins** choice for different unfoldings.

8.2.1.2 Side bins

Every migration bin can be built up from a linear combination of the **Signal bins** and thus inclusion of the **Migration bins** breaks the second paradigm and introduces ambiguity to the unfolding problem. With the addition of **Side bins** in the **Input** of unfolding,

[3] This subdivision is not pictured in figure 8.2 to improve its readability.

each being strongly correlated to its own **Migration bin**, the problem is removed. The **Side bin**, defined as a narrow slice outside of the cut value, controls the corresponding **Migration bin** and, in a way, allows the unfolding to directly measure the amount of migrations usually fully taken from the MC model.

For each **Side bin** two quality criteria are calculated. Diagram 8.3a helps to understand relevant definitions. In the example sketched in the figure a cut is introduced on the hadron level variable $x^{gen} < x_0^{gen}$. Likely, on the detector level the selection takes form $x^{rec} < x_0^{rec}$. The side bin is in this case defined as $x_0^{rec} < x^{rec} < x_1^{rec}$. Events both generated in the proper phase space and fulfilling the selection criteria are indicated by box **A**. Inclusion of **Migration bins** discussed in the previous section was motivated by events class **B**, migrating into the detector selection from outside of the hadron level phase space. Boxes **E** and **F** indicate events fulfilling the **Side bin** selection criteria. Events not reconstructed, but generated in the phase space (box **C**) are used for the efficiency correction. Finally, events both generated outside of the hadron level phase space and not reconstructed, labelled with box **D**, do not participate in the unfolding process.

The primary purpose of the **Side bin** is to provide the information about outside of the hadron level phase space, and thus the significance S defined as

$$S = E/(F + E) \tag{8.13}$$

should be kept reasonably high. However, to concentrate on the part of the generated phase space that is directly influencing the measurement, the relevance R defined as

$$R = B/(B + E) \tag{8.14}$$

should remain high. Both significance and relevance can be adjusted by changing the **Side bin** definition and should be treated more as a guideline than a systematic criteria of its determination. The typical $S(x_1^{rec})$ and $R(x_1^{rec})$ dependence is presented in figure 8.3b.

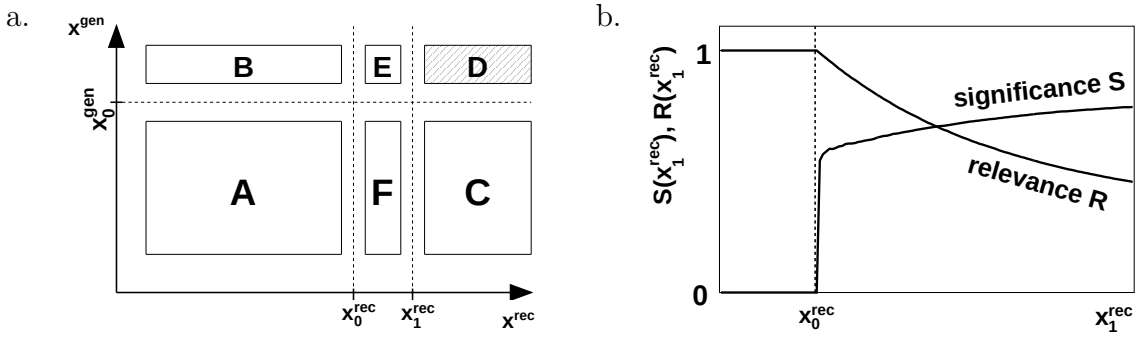

Figure 8.3: Classes of events for the example of a hadron level cut $x^{gen} < x_0^{gen}$ and detector level cut $x^{rec} < x_0^{rec}$. The side bin is defines as $x_0^{rec} < x^{rec} < x_1^{rec}$ (a). The typical behaviour of the significance and relevance as a function of the side bin definition (b).

Figure 8.4 shows the significance and relevance plots for all the **Side bins** used in this analysis corresponding to all[4] hadron level cuts. The vertical lines indicate the chosen x_1^{rec} cut values while the table 8.4 summarises the final choice of the side bin definitions. Since, as already mentioned, all the **Migration bins** are subbinned into $2E_T^\gamma \times 3\eta^\gamma$ bins, the corresponding **Side bins** are binned in a similar way.

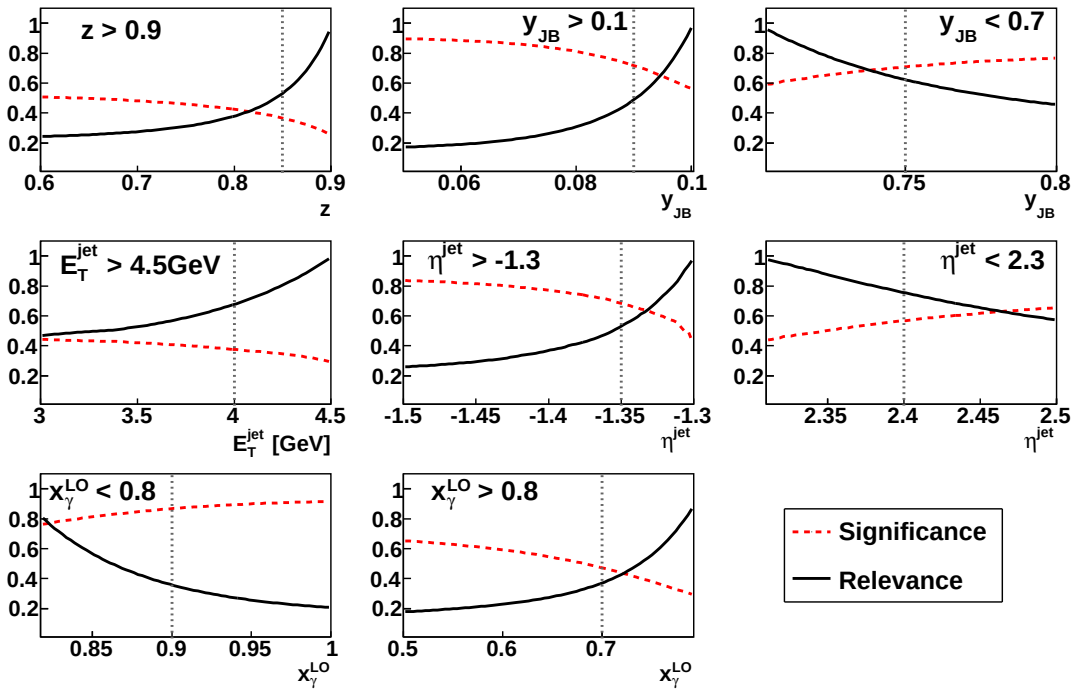

Figure 8.4: Significance and relevance of the **Side Bins** for all the hadron level cuts. The vertical lines indicate the x_1^{rec} values.

Hadron level cut definition	Side Bin definition
$z > 0.9$	$0.85 < z < 0.9$
$y > 0.1$	$0.09 < y < 0.1$
$y < 0.7$	$0.7 < y < 0.75$
$E_T^{jet} > 4.5$ GeV	$4.0 < E_T^{jet} < 4.5$ GeV
$\eta^{jet} > -1.3$	$-1.35 < \eta^{jet} < -1.3$
$\eta^{jet} < 2.3$	$2.3 < \eta^{jet} < 2.4$
$x_\gamma^{LO} < 0.8$	$0.8 < x_\gamma^{LO} < 0.9$
$x_\gamma^{LO} > 0.8$	$0.7 < x_\gamma^{LO} < 0.8$

Table 8.4: Definitions of the Side Bins.

[4] All but the cut on the Q^2 variable, which is handled separately.

8.2.1.3 Background bins

Every single **Reconstructed bin** and **Side bin** in real data contains an a priori unknown fraction of background (see chapter 6). Migration matrix definition includes one **Background bin** defined on the **Output** direction for each of the **Input bins**. In the example of figure 8.2, there are eight **Background bins** corresponding to six **Reconstructed Bins** and two **Side bins**. Due to special usage of the discriminator, there is a high correlation between **Background bins** and the lower discriminator bins of the corresponding **Input bins**. **Signal bins** and **Migration bins** couple to the high discriminator values, so adding **Background bins** does not introduce any ambiguity to the unfolding problem.

8.2.2 Selection efficiency correction

Events generated in the studied phase space, but not fulfilling the selection criteria are accommodated in a special underflow bin of the migration matrix. For every **Signal bin** the efficiency correction factor is defined as a ratio between the sum of weights of all the events generated in the particular **Signal bin** to the sum of weights of the events reconstructed either in one of the **Reconstructed bins** or in one of the **Side bins**. After the calculation of the output vector, the solution is appropriately corrected.

The trigger efficiency correction needs a special treatment, as the MC simulation differs in this respect significantly from the measured one in data. The correction factor $w^{Trigger}$, defined with the equation 4.4 is used to artificially decrease the simulated trigger efficiency. For each MC event the original event weight w is modified according to the formula: $w' = w \cdot w^{Trigger}$. Such modified w' weight is used to fill the migration matrix. In the same time, the corresponding underflow bin, responsible for the detector inefficiency is filled using the weight $w'' = w \cdot (1 - w^{Trigger})$. Since $w' + w'' = w$, and taking into account that the detector inefficiency bin does not participate directly in the unfolding procedure, the second paradigm holds and the unambiguity of the solution is preserved.

The final migration matrix **A** is normalised in a following way:

$$\sum_i A_{ij} = 1, \tag{8.15}$$

where the sum runs over every **Input bin** i and the independent normalisation is performed for every **Output bin** j.

8.2.3 Regularisation

The multidimensional output binning prevents the usage of the derivative or curvature regularisation schemes, as neighboring bins should not always remain connected. For example the highest E_T^γ bin stays next to the lowest E_T^γ bin of the next η^γ bin. Therefore, this analysis is using the regularisation on the size of the output. As explained earlier, the only applicable regularisation parameter τ choice method is the L-curve method (see

section 8.1.1). Figure 8.5 presents the curves obtained with the τ scanning for all the unfoldings performed to obtain the final results. The optimal choice of the working point is highlighted and the corresponding τ^2 value is printed.

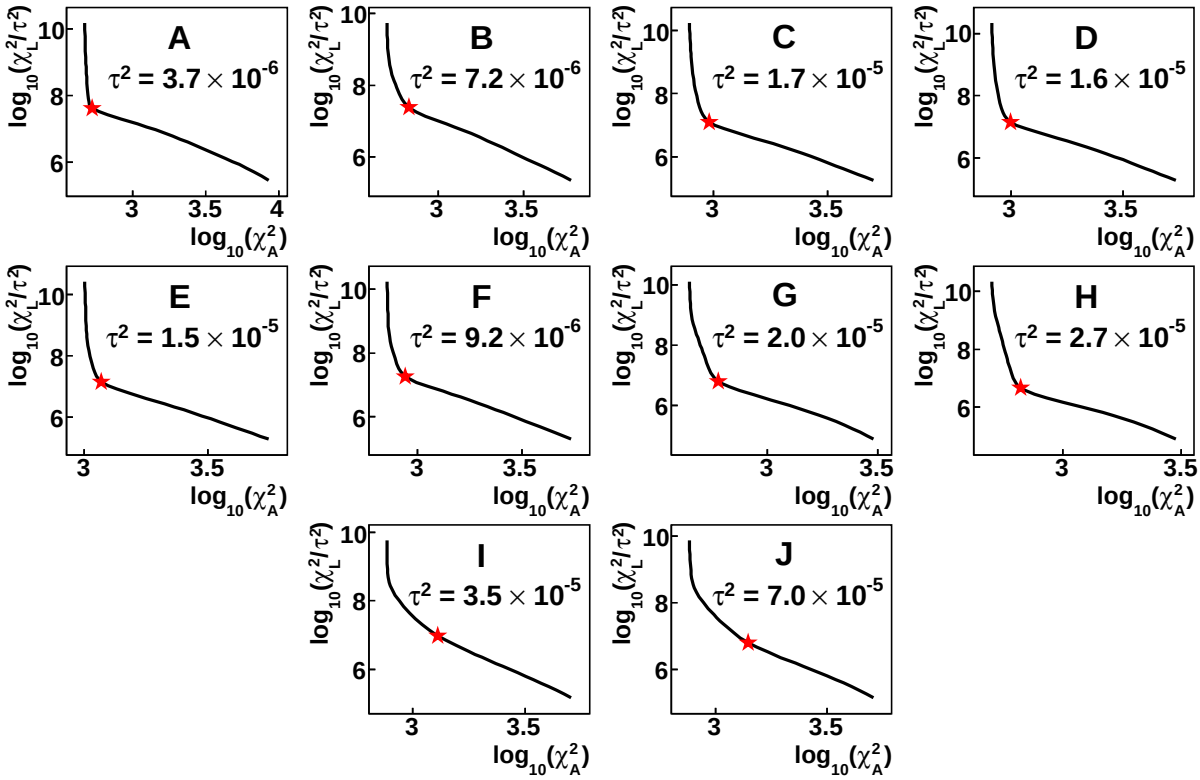

Figure 8.5: Scanning of the regularisation parameter τ for all the unfoldings performed in this analysis. The optimal choice of τ is indicated by the star.

8.2.4 Covariance matrix redefinition

The unfolding procedure is using the covariance matrix of the input bins $\mathbf{V}$. Usually the real covariance matrix is unknown but can be approximated by the measured errors of the input vector with the assumption of the statistical independence between bins:

$$\mathbf{V} = \begin{bmatrix} \epsilon_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_n \end{bmatrix}, \tag{8.16}$$

where ϵ_i is the measured error of the i^{th} bin in the input vector (usually square root of the sum of weights). In such an approach, if $\mathbf{V}$ is used for a χ^2 calculation, the statistical fluctuation between input bins may introduce a bias towards lower solutions, as has been shown in [124].

In order to illustrate the effect, one can consider the simple case with only two bins. The "Real" distribution is presented as a dashed red line in the figures 8.6a and 8.6b. Due to statistical reasons one of the events migrated from bin A into bin B (figure 8.6a) or from bin B into bin A (figure 8.6b). Both "Data" distributions are still statistically consistent with the original histogram. The errors on the "Data" histogram were approximated using number of observed events. Calculation of the χ^2_V function with one parameter a_v being the scale of the original histogram, leads to:

$$\chi^2_V(a_V) = \sum_i \frac{(N_{Data,i} - a_V \cdot N_{Real,i})^2}{\sigma^2_{Data,i}} \tag{8.17}$$

The minimisation of the χ^2_V in respect to the a_V is the usual way to determine the unfolded scale, but leads to systematically underestimated solutions (as one can see from the fitted a_V value is shown in figure 8.6). Correct results are obtained with usage of the real errors, taken directly from the "Real" distribution. In such a case one minimises

$$\chi^2_{VV}(a_{VV}) = \sum_i \frac{(N_{Data,i} - a_{VV} \cdot N_{Real,i})^2}{\sigma^2_{Real,i}} \tag{8.18}$$

and obtains the a_{VV} parameter as a solution. The same situation happens in case of simultaneous χ^2 minimisation of all four bins presented in the figure 8.6c.

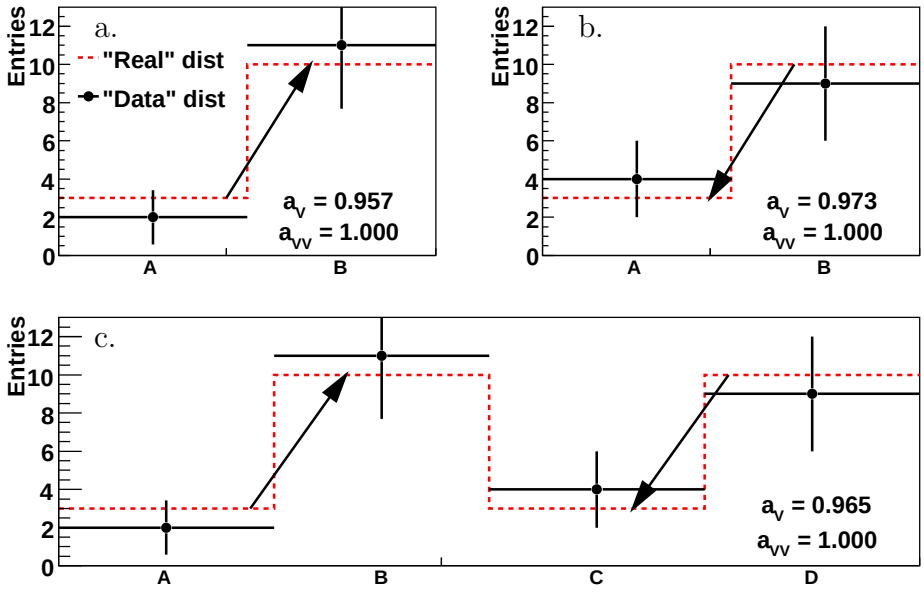

Figure 8.6: The example of two bin steep real distribution (dashed red) and an observed distribution with one event migration upstream the slope (a) or downstream the slope (b). The same for a double peak distribution with four bins (c).

In the case of this analysis similar effect can be observed by comparing the input vector $\vec{y}_{obs}$ with the vector of results folded back $\vec{y}_{folded} = \mathbf{A}\vec{x}_{true}$ for example with the help of

the *Pull* histogram filled with the $Pull_i$ contributions:

$$Pull_i = \frac{\vec{y}^{\,i}_{obs} - \vec{y}^{\,i}_{folded}}{\epsilon_i}. \tag{8.19}$$

where index i runs over all the bins of compared distributions. The *Pull* histogram should be a Gaussian type distribution with the mean value $Pull_0$ and standard deviation σ_{Pull}. A non zero $Pull_0$ indicates an overall bias and σ_{Pull} being greater than one signals a general incompatibility of the compared distributions. Figure 8.7a presents the *Pull* histogram as obtained when comparing $\vec{y}_{obs}$ and $\vec{y}_{folded}$ vectors for the example of **B** unfolding. The $Pull_0$ and σ_{Pull} parameters of the fitted Gaussian distribution are also printed.

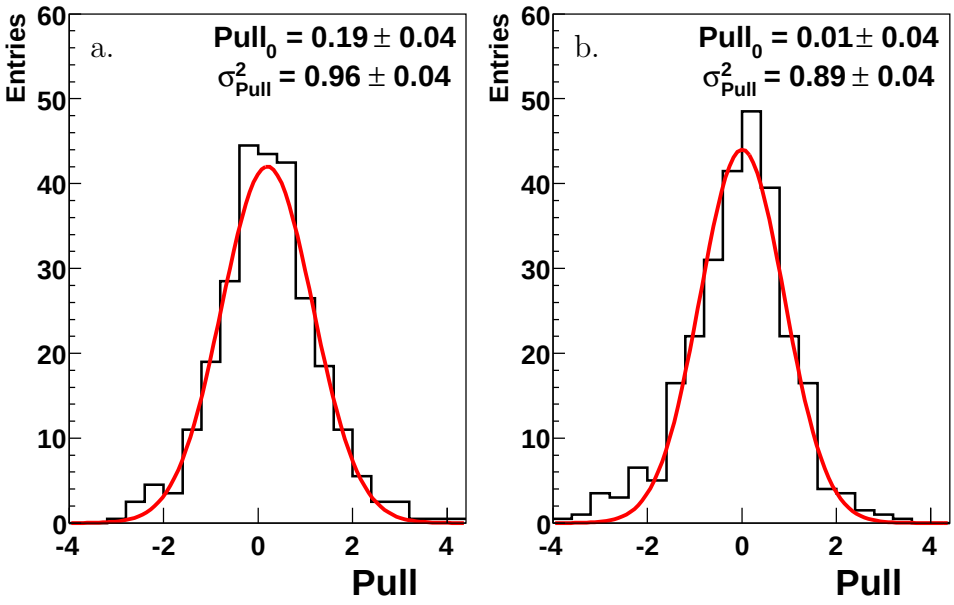

Figure 8.7: Pull histograms of the **B** unfolding after the first (a) and second (b) iteration.

As already noted in the simple two bins case, the bias can be eliminated by usage of real errors, or, if they are unknown, the expected ones. The unfolding procedure is iterated with the second iteration using redefined covariance matrix $\mathbf{V}'$:

$$\mathbf{V}' = \begin{bmatrix} \epsilon'_1 & 0 & \cdots & 0 \\ 0 & \epsilon'_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon'_n \end{bmatrix}, \tag{8.20}$$

where $\epsilon'_i = \epsilon_i \cdot \sqrt{(\mathbf{A}x_{true})^i / y^i_{obs}}$.

After the second iteration, the systematic bias is eliminated. One can see it for the example of **B** unfolding by comparing the plots in figure 8.7 (figure 8.7b presents *Pull* histogram after the second iteration). The table 8.5 lists the unfolding bias b defined as

$$b = \frac{|\vec{y}_{folded}| - |\vec{y}_{obs}|}{|\vec{y}_{obs}|} \times 100\% \tag{8.21}$$

as calculated after the first iteration and after the second iteration for each unfolding code. At the same time it presents the mean value $Pull_0$ and standard deviation σ^2_{Pull} of the fitted Gaussian distributions. The improvement after the second iteration is clearly visible, as both, the bias b and mean value of the fitted Gaussian function $Pull_0$ become consistent with zero. The σ^2_{Pull} is always below one indicating that both $\vec{y}_{obs}$ and $\vec{y}_{folded}$ are statistically consistent for each case. The higher number of iterations do not change the results.

Unfolding Code	b (1^{st})	b (2^{nd})	$Pull_0$ (1^{st})	$Pull_0$ (2^{nd})	σ^2_{Pull} (1^{st})	σ^2_{Pull} (2^{nd})
A	-4.9%	-1.4%	0.15 ± 0.04	-0.02 ± 0.04	0.99 ± 0.04	0.91 ± 0.04
B	-7.0%	-1.8%	0.19 ± 0.04	0.01 ± 0.04	0.97 ± 0.04	0.89 ± 0.03
C	-8.8%	-1.0%	0.20 ± 0.02	0.00 ± 0.02	0.78 ± 0.02	0.74 ± 0.02
D	-8.4%	-0.7%	0.18 ± 0.02	-0.04 ± 0.02	0.78 ± 0.02	0.78 ± 0.02
E	-8.5%	$+0.4\%$	0.19 ± 0.02	-0.03 ± 0.03	0.82 ± 0.02	0.79 ± 0.02
F	-7.6%	-0.9%	0.17 ± 0.03	-0.04 ± 0.03	0.81 ± 0.02	0.83 ± 0.03
G	-0.5%	$+5.5\%$	0.17 ± 0.02	0.02 ± 0.02	0.66 ± 0.02	0.67 ± 0.02
H	-5.5%	$+4.2\%$	0.18 ± 0.02	0.00 ± 0.02	0.69 ± 0.02	0.68 ± 0.03
I	-11.8%	-0.9%	0.20 ± 0.02	-0.04 ± 0.03	0.75 ± 0.02	0.78 ± 0.02
J	-14.3%	-2.5%	0.24 ± 0.02	0.00 ± 0.02	0.80 ± 0.02	0.76 ± 0.02

Table 8.5: The unfolding bias b and Gaussian distributions parameters fitted to the $Pull$ histograms after the first (1^{st}) and second (2^{nd}) unfolding iteration.

8.2.5 Q^2 selection correction

The cross sections presented in this thesis are quoted in the photoproduction kinematic region with four-momentum transfer squared $Q^2 < 1\,\mathrm{GeV}^2$. This photoproduction definition was used for H1 measurements with HERA I data. The QCD calculation produced in order to be compared to data are generated in the same photoproduction region. During HERA II upgrade, the SPACAL calorimeter acceptance was decreased and it provides the electron veto for $Q^2 > 3\,\mathrm{GeV}^2$. Since the Q^2 selection relies mainly on the electron veto in the SPACAL, the results need to be correctly corrected for the existing Q^2 gap using special MC study.

The unfolding input vector $\vec{y}_{obs}$ was filled without any direct condition on the reconstructed Q^2. The DIS background has been minimised by the scattered electron veto and y condition (see section 5.1). For each input vector two additional histograms were filled with MC events: $\vec{y}_{obs,MC}$ and $\vec{y}_{obs,MC,Q2}$. Vector $\vec{y}_{obs,MC}$ was filled with exactly the same selection conditions as in the case of $\vec{y}_{obs}$, while in the case of $\vec{y}_{obs,MC,Q2}$ additional request on generated four momentum transfer squared $Q^2 < 1\,\mathrm{GeV}^2$ was applied. The

contribution of higher Q^2 values was subtracted by correcting each bin i of data input vector:

$$y'^{\,i}_{obs} = y^i_{obs} \times \frac{y^i_{obs,MC,Q2}}{y^i_{obs,MC}} \tag{8.22}$$

The correction affects mainly cross section in the backward region with its maximum value of 5%. The ratio $y_{obs,MC,Q2}/y_{obs,MC}$ is presented in figure 8.8 as a function of pseudorapidity of the photon candidate.The final results are correctly quoted in the photoproduction range of $Q^2 < 1\ GeV^2$.

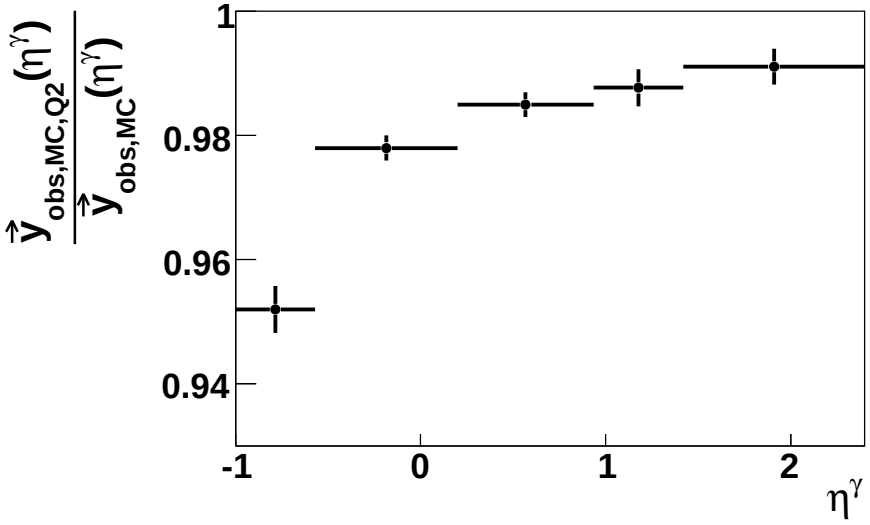

Figure 8.8: The Q^2 selection correction factor as a function of pseudorapidity of the photon candidate.

8.3 Error matrix evaluation

The final error matrix $\mathbf{E_x}$ consists of three components and is built in the following way:

$$\mathbf{E_x} = \mathbf{E_u} + \mathbf{E_{mc}} + \sum_i \mathbf{E^i_{sys}} \tag{8.23}$$

where $\mathbf{E_u}$ is the covariance matrix of the unfolding output bins calculated with the equation 8.12, $\mathbf{E_{mc}}$ is the error originating from the limited statistics of MC used to determine the migration matrix and $\mathbf{E^i_{sys}}$ is the error due to particular systematic effects i. Both $\mathbf{E_{mc}}$ and $\mathbf{E_{sys}}$ are calculated using error propagation rules [99] shortly explained in this section.

Considering n_η functions η_i, building vector $\vec{\eta}$ and depending on n_θ variables θ: $\eta_i(\theta_1, \cdots \theta_{n\theta})$ the covariance matrix $U_{ij} = cov|\eta_i, \eta_j|$ can be approximated by the first order in the Taylor

expansion series:

$$\mathbf{U}_{i,j} = \sum_{k,l} \frac{\partial \eta_i}{\partial \theta_k} \frac{\partial \eta_j}{\partial \theta_l} \Bigg|_{\hat{\theta}} \mathbf{V}^{\theta}_{\mathbf{k,l}} \tag{8.24}$$

with $\hat{\theta}$ being the estimator of θ and $\mathbf{V}^{\theta}_{\mathbf{k,l}} = cov|\theta_k, \theta_l|$. In the case of this analysis $\vec{\eta} = \vec{x}_{true}$, while $m \times n$ $\mathbf{A}_{i,j}$ elements of migration matrix $\mathbf{A}^{(m \times n)}$ are used as variables θ. In such a case, using the equation 8.11 one can calculate derivatives:

$$\frac{\partial \eta_i}{\partial \theta_k} \equiv \frac{\partial x^i_{true}}{\partial A_{p,q}} = (\mathbf{E_u})_{i,p} \times (\mathbf{V}(\vec{y}_{obs} - \mathbf{A}\vec{x}_{true}))_q - (\mathbf{E_u A^T V})_{i,q} \times (\vec{x}_{true})_p, \tag{8.25}$$

where indices i and p run over unfolding output bins and q runs over input bins.

The covariance matrix $\mathbf{U}$ takes then the form:

$$\mathbf{U}_{i,j} = \sum_{p,q} \sum_{r,s} \frac{\partial x^i_{true}}{\partial A_{p,q}} \frac{\partial x^j_{true}}{\partial A_{r,s}} \Bigg|_{\hat{A}} \mathbf{V}^{\mathbf{A}}_{\mathbf{(p,q),(r,s)}}, \tag{8.26}$$

where indices i, j, p and r run over unfolding output bins, while q and s are input bin indices.

8.3.1 MC statistics error propagation

The migration matrix as used in this analysis needs to be sufficiently populated to avoid large statistical fluctuations in the determination of the matrix itself. Special care was taken to include the MC statistical fluctuations in the final errors. The error originating from the limited statistics of the MC is calculated with each migration matrix element treated as independent source of uncorrelated error. In such a case the covariance matrix $\mathbf{V}^{\mathbf{A}}_{\mathbf{(p,q),(r,s)}}$ takes the form

$$\mathbf{V}^{\mathbf{A}}_{\mathbf{(p,q),(r,s)}} = \begin{cases} 0 & \text{if } (p,q) \neq (r,s) \\ \sigma^2_{p,q} & \text{if } (p,q) = (r,s) \end{cases} \tag{8.27}$$

where $\sigma^2_{p,q}$ is the statistical error of the element $\mathbf{A_{p,q}}$. The equation 8.26 simplifies to

$$\mathbf{U}_{i,j} = \sum_{p,q} \frac{\partial \eta_i}{\partial A_{p,q}} \frac{\partial \eta_j}{\partial A_{p,q}} \Bigg|_{\hat{A}} \sigma^{\mathbf{2}}_{\mathbf{p,q}} \equiv \mathbf{E_{mc}} \tag{8.28}$$

In practice, since migration matrix is normalised, two errors are propagated. One coming from the matrix binomial error and the second, originating from the limited statistics of the normalisation factor.

The MC statistical error contributes to at most 10% of the final error in the low η^{γ}-high E^{γ}_T region. Typically it is consistently below 5% of the final error and thus considered to be reduced to negligible level over the whole studied phase space.

8.3.2 Systematic error propagation

The systematic error has been incorporated into the same error propagation formalism. Systematic shifts, applied to the MC used for determination of unfolding matrix $\mathbf{A}$ can be directly translated into the error $\sigma^2_{p,q}$ associated with each of the unfolding matrix elements $A_{p,q}$. This error can now be accordingly propagated to the final results. Unlike the previous case though, for a given systematic shift all the errors are propagated in a correlated way. In this case covariance matrix takes the form:

$$\mathbf{V}^{\mathbf{A}}_{(\mathbf{p},\mathbf{q}),(\mathbf{r},\mathbf{s})} = \sqrt{\sigma^2_{p,q}}\sqrt{\sigma^2_{r,s}} \tag{8.29}$$

and equation 8.26:

$$\mathbf{U}_{i,j} = \sum_{p,q}\left(\left.\frac{\partial x^i_{true}}{\partial A_{p,q}}\right|_{\hat{A}}\sqrt{\sigma^2_{p,q}}\right) \times \sum_{r,s}\left(\left.\frac{\partial x^j_{true}}{\partial A_{r,s}}\right|_{\hat{A}}\sqrt{\sigma^2_{r,s}}\right) \equiv \mathbf{E_{sys}}. \tag{8.30}$$

The following systematic error sources were studied:

- **Trigger efficiency determination**
  The trigger efficiency determination is explained in the section 4.3.2. The alternative set of monitor triggers are used to evaluate alternative trigger correction and the alternative migration matrix $\mathbf{A}'$. The difference between both matrices is taken as a uncertainty of the matrix elements and propagated to the final results. The systematic error contribution due to trigger efficiency determination varies between 1% and 3%.

- **Luminosity uncertainty**
  The luminosity for the studied period is known with the precision of 3.4%. This systematic error is added to the final error matrix in a fully correlated way.

- **Track Efficiency**
  The track finding efficiency in the CJC and CIP enters the analysis due to the veto applied to the photon candidates. It is evaluated by performing the analysis with the pure CIP veto and with pure CJC veto. The difference between results is at most 2.5% and such systematic uncertainty is propagated to the final results. Since the results were calculated using the combined CIP and CJC veto, such an error might be slightly overestimated.

- **DIS background and Q^2 correction**
  Due to high reliance on the MC for the Q^2 correction described in section 8.2.5, the full correction is taken as a systematic error. It leads to 3% uncertainty on the final results on average, with 5% particularly for backwards photons.

- **Cluster shower shapes description**
  The cluster shapes MC description study is presented in the section 7.1. The alternative migration matrix is build using MC with cluster shapes stretched within the

evaluated uncertainties. The variables R_T, HCF and $HCellF$ are treated simultaneously, as one may expect that change of the calorimeter shower simulation would affect all three of them at the same time. The final error due to the propagation of their uncertainties becomes the leading error due to the fact that small changes in the width of the cluster may shift events from the **Signal** unfolding output to the **Background** part and vice versa. The error on the total inclusive cross section has been evaluated as 11%. For a single differential cross section the error varies between 10% up to 25% in the forward region of the detector. S_T, K_T and FLF uncertainties are propagated independently and their influence is much lower (1 - 3%).

- **Photon energy scale**
  The photon energy scale is studied in section 7.2. The alternative migration matrix is evaluated with 1% correction to the MC photon energy correction for photons with pseudorapidity $\eta^\gamma < 1.4$ and 4% for photons with $\eta^\gamma > 1.4$, resulting in the final error of roughly 1.5% on average.

- **Hadronic energy scale**
  The study of the MC description of hadronic system energy is presented in section 7.3. A 2% correction is applied to the energy of all hadronic particle candidates in the MC and alternative migration matrix is evaluated. The correction affects the efficiency loss due to y cut for all cross sections and E_T^{jet} cut for the exclusive selection. The error associated with this effect is below 1%.

- **The polar angle of the photon**
  The MC description of the polar angle resolution has been evaluated as 3 mrad [43]. For the pseudorapidity range $\eta^\gamma > 1.4$ it is increased to 4 mrad. The propagated error caused by this uncertainty is included in the final errors, but is of negligible level.

- **Direct to resolved event class ratio**
  The relative scales of the direct and resolved prompt photon MC production have been evaluated in section 7.4 together with their uncertainties. Due to lower hadronic activity for the direct cases one may expect differences in the acceptance corrections, mostly due to the photon isolation criteria. The calculated error is usually of 1% order.

- **Detector dead material simulation**
  An uncertainty in the description of the dead material in the simulation is accounted for by varying the probability of photon conversion before the calorimeter by 10%. For polar angles $\theta < 20^\circ$ it is varied by 30% because of more dead material in the forward region of the detector. This results in a 1% error in the central region and 3% in the most forward η^γ bin.

All the evaluated systematic error matrices were added and included in the final error matrix $\mathbf{E_x}$. Due to the high sensitivity of the signal extraction to the cluster shape description, it is by far most significant source of error.

8.4 Cross section building

In the previous section the details of the unfolding procedure were introduced. As the result of all the steps the multidimensional unfolding output and error matrix are obtained. In this part, the construction of final cross section is explained.

8.4.1 Bin averaging procedure

The output of the unfolding procedure is the multidimensional vector $\vec{x}_{true}$ and corresponding error matrix $\mathbf{E_x}$. In most cases final results are quoted as single differential cross sections, where unfolding output was projected along all unused dimensions. Such an approach allows the reduction of the final error, particularly in case of averaging over likely produced during unfolding anticorrelated bins. Table 8.6 presents the averaged dimensions for all the final cross sections.

Cross section	Variable	Unfolding Code	Output Variables	Averaging
i	E_T^γ	**A**	$E_T^\gamma \times \eta^\gamma$	η^γ
ii	η^γ	**A**	$E_T^\gamma \times \eta^\gamma$	E_T^γ
iii	E_T^γ	**B**	$E_T^\gamma \times \eta^\gamma$	η^γ
iv	η^γ	**B**	$E_T^\gamma \times \eta^\gamma$	E_T^γ
v	E_T^{jet}	**C**	$E_T^\gamma \times \eta^\gamma \times E_T^{jet}$	E_T^γ, η^γ
vi	η^{jet}	**D**	$E_T^\gamma \times \eta^\gamma \times \eta^{jet}$	E_T^γ, η^γ
vii	x_γ^{LO}	**E**	$E_T^\gamma \times \eta^\gamma \times x_\gamma^{LO}$	E_T^γ, η^γ
viii	x_p^{LO}	**F**	$E_T^\gamma \times \eta^\gamma \times x_p^{LO}$	E_T^γ, η^γ
ix	$p_\perp$	**G**	$E_T^\gamma \times \eta^\gamma \times p_\perp$	E_T^γ, η^γ
x	$\Delta\phi$	**H**	$E_T^\gamma \times \eta^\gamma \times \Delta\phi$	E_T^γ, η^γ
xi	$p_\perp$	**I**	$E_T^\gamma \times \eta^\gamma \times p_\perp$	E_T^γ, η^γ
xii	$\Delta\phi$	**J**	$E_T^\gamma \times \eta^\gamma \times \Delta\phi$	E_T^γ, η^γ

Table 8.6: Averaging dimensions for all the final cross sections.

The number of photons in the final bin k with all contributing $\vec{x}_{true}$ bins i is calculated as

$$N_k = \sum_i (\vec{x}_{true})_i \tag{8.31}$$

The error matrix $\mathbf{E}$ is obtained in a similar way

$$\mathbf{E}_{k,l} = \sum_{i,j} \mathbf{E_x}_{i,j} \tag{8.32}$$

where bins i are output bins contributing to the final bin k and bins j are the ones contributing to l. Since the elements of matrix $\mathbf{E_x}$ can take positive as well as negative values, error cancellation is possible (though regularisation largely decrease its probability).

8.4.2 Error matrix interpretation

The unfolding procedure allows the determination of the error matrix that includes all the correlations between any of the output bins. The averaging procedure translates those correlations into correlations between final cross section bins. In this section, the splitting of the final error matrix into an easy to visualise and interpret fully correlated part and fully uncorrelated parts is explained.

The error matrix, in case of n fully correlated bins takes the form of

$$\mathbf{E_a} = \begin{bmatrix} a_1a_1 & a_1a_2 & \cdots & a_1a_n \\ a_2a_1 & a_2a_2 & \cdots & a_2a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1 & a_na_2 & \cdots & a_na_n \end{bmatrix} \tag{8.33}$$

with n independent parameters a_i. In the same time, the error matrix for n fully uncorrelated bins can be described by n parameters δ_i:

$$\mathbf{E}_\delta = \begin{bmatrix} \delta_1^2 & 0 & \cdots & 0 \\ 0 & \delta_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_n^2 \end{bmatrix}. \tag{8.34}$$

The assumption that the measured error matrix is a linear sum of the fully uncorrelated part and the fully correlated part leads to the equation:

$$\mathbf{E} = \mathbf{E_a} + \mathbf{E}_\delta. \tag{8.35}$$

For $e_{i,j}$ being the elements of the error matrix $\mathbf{E}$, one can produce $n(n-1)/2$ equations with $2n$ parameters.

$$e_{i,i} = a_i^2 + \delta_i^2, \tag{8.36}$$

$$e_{i,j} = a_ia_j \qquad \text{for } i \neq j. \tag{8.37}$$

Since the number of parameters exceeds number of equations, only an approximate solution can be found. If one in addition eliminate n parameters, by requesting the strict hold of the equations 8.36, the remaining parameters a_i can be determined by the minimisation of $n-$dimensional χ^2 function

$$\chi^2(a_1...a_n) = \sum_{i,j,i\neq j} \frac{(a_ia_j - e_{ij})^2}{e_{ii}e_{jj}}, \tag{8.38}$$

while

$$\delta_i = \sqrt{e_{i,i} - a_i^2}. \tag{8.39}$$

Such obtained parameters δ_i can in principle take genuinely complex values, as there are no limits forcing minimisation to produce a_i^2 smaller than $e_{i,i}$. Since this is not the case though, it gives additional confidence in the validity of the procedure. The minimisation of χ^2 function leads to a_i parameters reproducing e_{ij} elements mostly within few percentage precision.

8.4.3 Cross sections

The output of the bin averaging procedure explained in section 8.4.1 is the final number of prompt photons N_i measured in bin i. The error treatment associates with it a fully correlated error a_i and a fully uncorrelated error δ_i. The single differential cross section in bin i of variable x and double differential cross section in bin i of variable x and bin j of variable y is calculated with the formula

$$(\frac{d\sigma}{dx})_i = \frac{N_i}{L \cdot w_{x,i}} \pm \frac{a_i}{L \cdot w_{x,i}}(corr.) \pm \frac{\delta_i}{L \cdot w_{x,i}}(uncr.) \quad (8.40)$$

$$(\frac{d\sigma}{dxdy})_{i,j} = \frac{N_{i,j}}{L \cdot w_{x,i} \cdot w_{y,j}} \pm \frac{a_{i,j}}{L \cdot w_{x,i} \cdot w_{y,j}}(corr.) \pm \frac{\delta_{i,j}}{L \cdot w_{x,i} \cdot w_{y,j}}(uncr.) \quad (8.41)$$

$$(8.42)$$

where L is the integrated luminosity (see section 3.1) of the considered event selection (340 pb^{-1}), and $w_{x,i}$ ($w_{y,j}$) is the width of i^{th} (j^{th}) bin of variable x (y).

The total cross section σ_{tot} is determined for each unfolding by averaging over all the **Signal Bins** of $\vec{x}_{true}$ within the phase space. Since it is a single number, and the error division into a correlated and uncorrelated error does not make sense, the total cross sections are quoted together with statistical error (averaged over the sum of $\mathbf{E_{unf}}$ and $\mathbf{E_{mc}}$ and systematic error (averaged over the sum of $\mathbf{E^i_{sys}}$).

8.5 Total cross section consistency check

The unfolding procedure has been repeated ten times, in order to calculate cross sections binned in different variables. Particularly, one of those unfoldings (**A**) leads to the inclusive prompt photon production cross sections, five (**B, C, D, E, F**) were performed in the exclusive photon plus jet phase space, two (**G, H**) in the direct enhanced phase space and two (**I, J**) in the resolved enhanced phase space. Different unfoldings in the same phase space produce the same cross sections projected on a different direction, so the comparison of the total cross sections can be treated as a consistency check.

Figure 8.9 presents the comparison of the total cross sections in the exclusive phase space obtained with five different unfoldings. The inner error bars represent the statistical error while outer error bars statistical and systematic added in quadrature. The shaded area and horizontal lines indicate error ranges for the most precise and least complex unfolding **B**. The consistency of all obtained total cross sections increase the confidence level in the relatively complex method explained in this chapter.

8.6 Toy Monte Carlo study

The most common way of checking the quality of the unfolding procedure is a toy Monte Carlo study like the one presented in this section. The whole set of toy MC samples

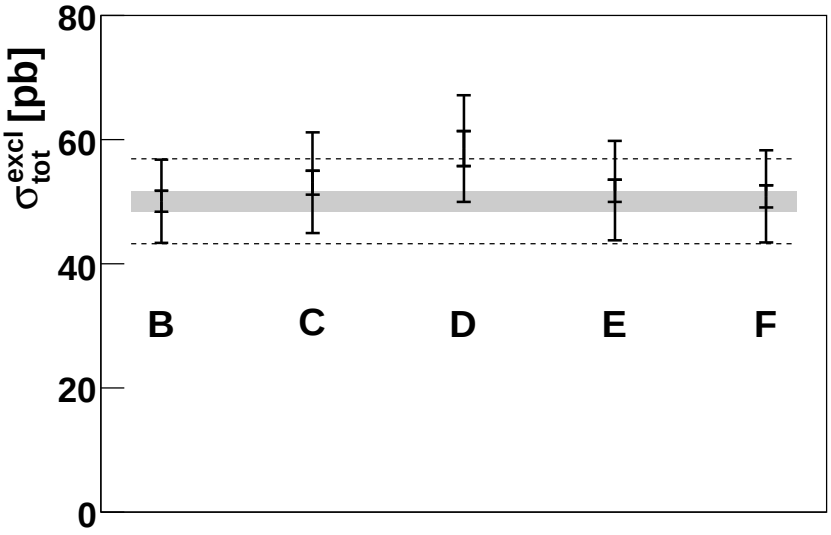

Figure 8.9: The exclusive photon plus jet total cross sections obtained with five different unfoldings. The inner error bars represent the statistical error, outer error bars statistical and systematic added in quadrature. The shaded area and horizontal lines indicate error ranges for the unfolding **B**.

is subject to the unfolding procedure and the results are compared to the known MC distribution. Three sets of toy MC samples were produced for the purpose of this analysis by reweighting the original MC set I:

- The set of seven reweighting functions $f^x_{E_T}$ designed to reweight the MC E^γ_T distribution in such a way that the reweight value for events with $E^\gamma_T = 6\,\mathrm{GeV}$ is equal to one, while events with $E^\gamma_T = 15\,\mathrm{GeV}$ are weighted up by x percentage:

$$f^x_{E_T} = x \times E^\gamma_T/900\,\mathrm{GeV} - x/150 - 1, \tag{8.43}$$

  with the chosen x values are taken from the set $5, 10, 20, 50, 100, 150, 250$.

- The set of seven reweighting functions f^x_η designed to reweight the MC η^γ distribution in such a way that the reweight value for events with $\eta^\gamma = -1$ is equal to one, while events with $\eta^\gamma = 2.4$ are weighted up by x percentage:

$$f^x_\eta = x \times \eta^\gamma/340 + x/340 + 1, \tag{8.44}$$

  with the chosen x values are taken from the set $5, 10, 20, 50, 100, 150, 250$.

- The double dimensional $E^\gamma_T - \eta^\gamma$ reweighting function $f^{toy}_{E_T,\eta}$ produced to test the unfolding in the extreme case. The function implements reweighting in η^γ corresponding to f^{450}_η convoluted with non linear reweighting in E^γ_T of the scale ranging from $f^{350}_{E_T}$ up to even $f^{1100}_{E_T}$:

$$f^{toy}_{E_T,\eta} = (E^\gamma_T/GeV)^2/40 + \eta^\gamma/2 \tag{8.45}$$

The double dimensional $E_T^\gamma - \eta^\gamma$ distribution of original MC set I and a $f^{toy}_{E_T,\eta}$ reweighted MC is presented in the figure 8.10.

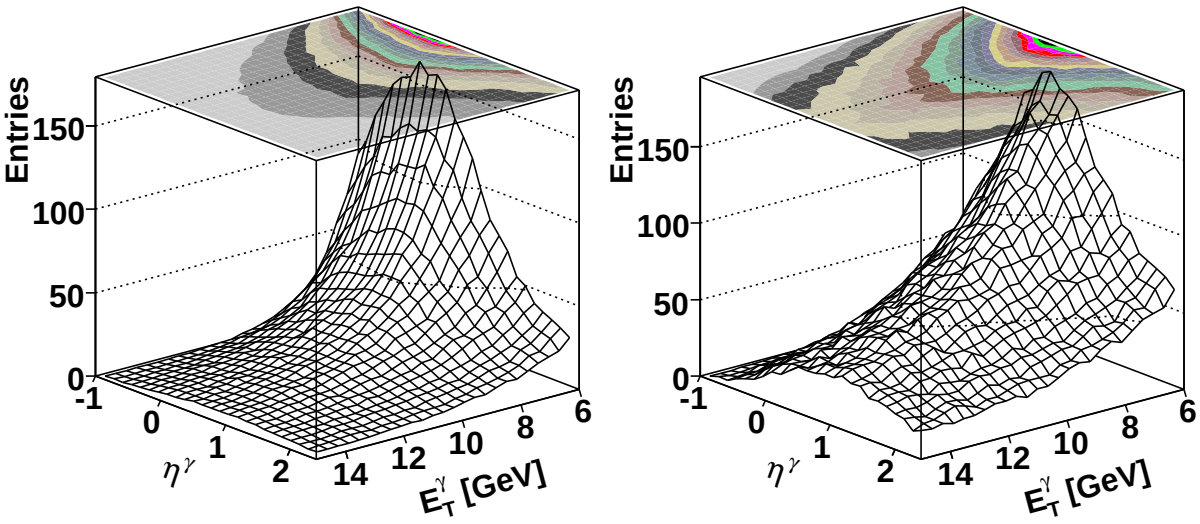

Figure 8.10: The E_T^γ-η^γ distribution of the original MC set I and a MC reweighted with the $f^{toy}_{E_T,\eta}$ function.

8.6.1 Fitting procedure

The quality of the whole unfolding procedure is additionally verified by using an alternative method of cross section determination. The method explained below was used to determine the prompt photon cross sections in most of the prompt photon publications based on the HERA data (i.e. [43], [35]) combines the discriminator fitting with bin-to-bin acceptance correction.

In this method, the discriminator D (see section 6.3) is binned in the transverse energy E_T^γ and pseudorapidity η^γ of the photon in bins directly corresponding to the binning used for the output. In the case of this analysis, that corresponds to twenty bins of $4(E^\gamma_{T,OUT4}) \times 5(\eta^\gamma_{OUT5})$ grid with the actual bin edges listed in appendix table A-4. A similar grid is filled using pure signal MC and pure background MC. In every bin i the actual amount of prompt photon signal is extracted using independent minimal-χ^2 fit of the normalised signal (d_{sig}) and background (d_{bkg}) discriminator distributions to the data. The χ^2_D function is defined as

$$\chi^2_D(N_{sig,i}, N_{bkg,i}) = \sum_j \frac{(N_{data,i,j} - N_{bkg,i}d_{bkg,j} - N_{sig,i}d_{sig,j})^2}{\sigma^2_{data,i,j} + N^2_{bkg,i}\sigma^2_{bkg,j} + N^2_{sig,i}\sigma^2_{sig,j}} \tag{8.46}$$

where σ_{data}, σ_{sig} and σ_{bkg} are the data, signal and background errors, N_{sig} and N_{bkg} are parameters representing fitted number of signal and background events and the sum runs

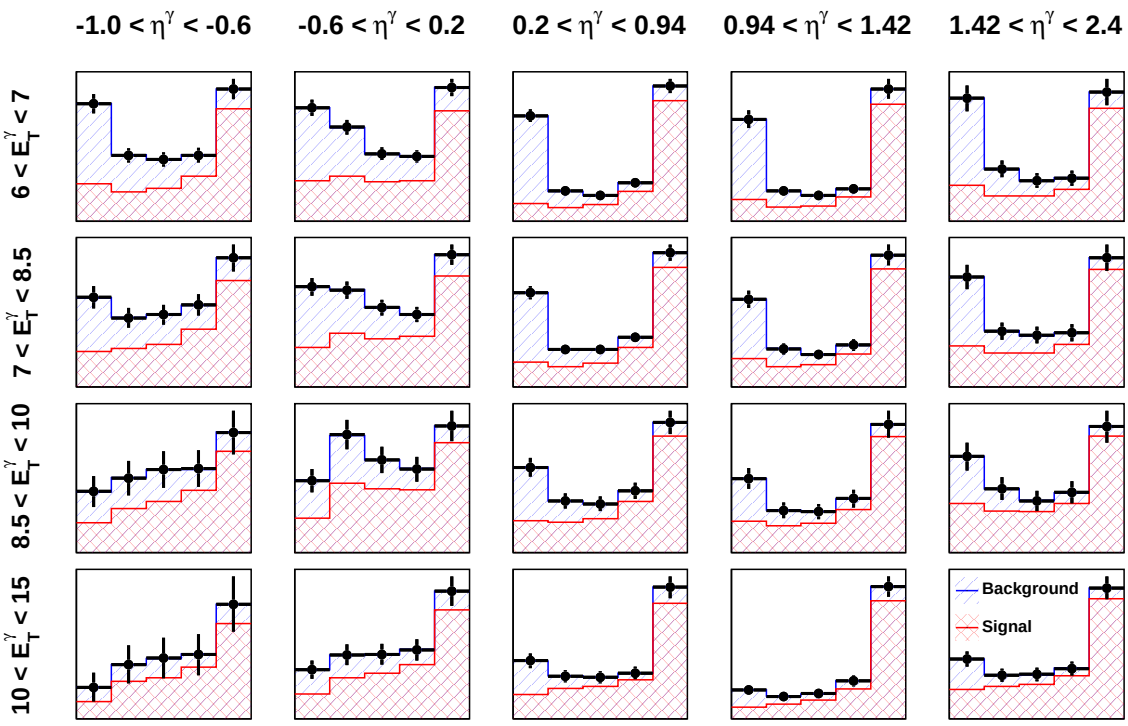

Figure 8.11: The discriminator fit performed in bins of E_T^γ and η^γ for one of the toy Monte Carlos.

over all the bins j of the discriminator distribution. The example of the fit in all twenty analysis bins is shown in figure 8.11.

The fitted number of signal events is corrected for the acceptance effects with the acceptance correction a defined for bin i as

$$a_i = \frac{N_{rec,i}}{N_{gen,i}} \tag{8.47}$$

where $N_{gen,i}$ is the number of prompt photons generated in bin i and $N_{rec,i}$ is the number of reconstructed signal events in the same bin. Acceptance correction factor defined in such a way incorporates the detector acceptance and efficiency effects as well as migrations between bins.

The final single differential cross sections in bins of variable x in the bin i is then determined using formula

$$(\frac{d\sigma}{dx})_i = \frac{1}{L \cdot w_{x,i}} \sum_j \frac{N_{sig,j}}{a_j} \tag{8.48}$$

where L is the integrated luminosity, $w_{x,i}$ is the width of i^{th} bin of variable x and the index j runs over all the contributing bins. More detailed explanation of the method is given in [66].

8.6.2 Toy MC results

The background contamination of real data was simulated by mixing all the toy MC samples with background MC events (MC set XVI). The statistics of the toy MC was chosen such as to simulate the expected data statistics. The E_T^γ and η^γ distributions have been unfolded using unfolding matrix filled with the original MC. Since the distributions of the unfolded MC and MC used to fill the matrix differs, the model inaccuracy is artificially introduced in a controlled way. The same distributions were determined using the fitting procedure and bin-to-bin correction where original MC was used to fill signal and background discriminator distributions as well as to determine the acceptance corrections.

In order to compare results to the real MC distribution the χ^2 variable is calculated according to the formula:

$$\chi^2 = \sum_i \frac{(R_i - T_i)^2}{\delta R_i}, \tag{8.49}$$

where T_i is the toy MC cross section in the bin i, R_i is the cross section obtained by unfolding or fitting procedure in bin i, δR_i is the calculated error of the results in bin i and index i runs over all the E_T^γ or η^γ bins. All the calculated χ^2 results are presented in table 8.7.

Toy MC	Fitting χ^2/ndf		Unfolding χ^2/ndf	
sample	E_T^γ	η^γ	E_T^γ	η^γ
$f_{E_T}^{005}$	0.143	0.190	0.075	0.110
$f_{E_T}^{010}$	0.155	0.198	0.073	0.106
$f_{E_T}^{020}$	0.175	0.214	0.070	0.096
$f_{E_T}^{050}$	0.245	0.270	0.083	0.068
$f_{E_T}^{100}$	0.380	0.388	0.143	0.034
$f_{E_T}^{150}$	0.528	0.530	0.240	0.014
$f_{E_T}^{250}$	0.343	0.412	0.080	0.198
f_{η}^{005}	0.138	0.186	0.078	0.118
f_{η}^{010}	0.140	0.190	0.078	0.118
f_{η}^{020}	0.150	0.198	0.078	0.120
f_{η}^{050}	0.175	0.222	0.078	0.128
f_{η}^{100}	0.215	0.264	0.078	0.142
f_{η}^{150}	0.258	0.312	0.078	0.160
f_{η}^{250}	0.343	0.412	0.080	0.198
$f_{E_T,\eta}^{toy}$	3.70	3.32	1.43	0.120

Table 8.7: The difference between true E_T^γ and η^γ distributions of toy MC and distributions obtained by fitting or unfolding procedures quantified by χ^2/ndf values.

Since in almost all the cases the calculated χ^2/ndf is strongly below one[5], conclusion can be drawn that both methods extract the correct cross sections within their own determined uncertainties. As expected, unfolding produces consistently lower values of χ^2, which points to it as a more reliable method. In addition, the impact of model inaccuracy can be observed, as χ^2 steadily rises with the reweight factor. The importance of proper model description seems to be more important for E_T^γ, as χ^2 rises faster. The reason for that are much lower bin purities in case of E_T^γ bins compared to η^γ case.

In figure 8.12 the results of the toy MC study with $f^{toy}_{E_T,\eta}$ reweighting scheme are presented. One can see that given the amount of the distortion applied to the toy MC, both unfolding and fitting procedures fail to correctly determine the E_T^γ distribution, with unfolding being more correct in the lowest E_T^γ bin. In case of η^γ, unfolding manages to determine the correct distribution, while fitting fails to do so. In addition, the correlation between bins, taken into account in the unfolding procedure leads to an increase of the evaluated error.

One may observe that fitting leads to systematically higher cross sections, which may be explained by the harder E_T^γ slope in the model used to extract the cross sections (see figure 8.10). That leads to the overestimation of migration from below $E_T^\gamma = 6\,\text{GeV}$ cut value. The unfolding better deals with this problem (see χ^2/ndf values in table 8.7).

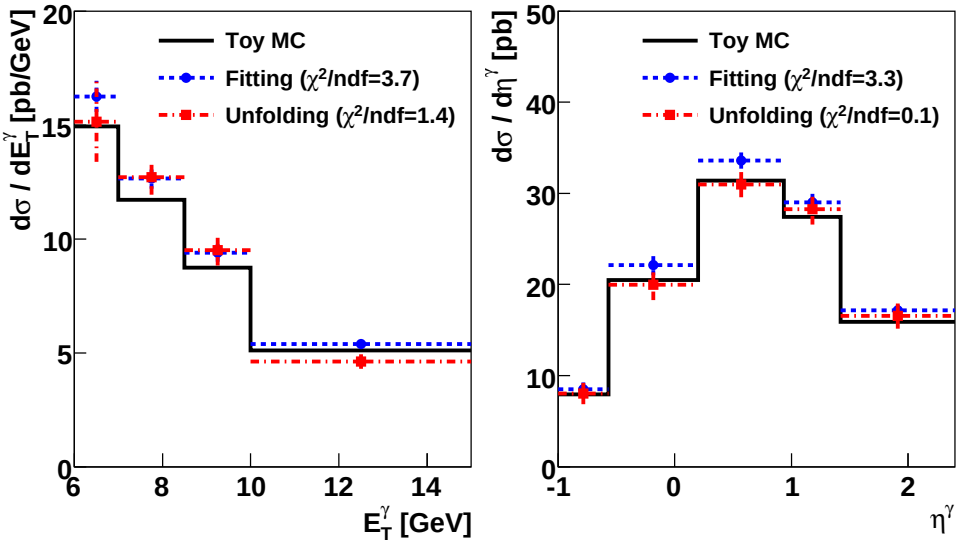

Figure 8.12: The E_T^γ and η^γ distribution of the toy Mc compared to the results obtained with the unfolding and fitting procedures.

The bin-to-bin correction is known to produce unreliable results in case of large model inaccuracies, while unfolding is expected to produce correct results even with the particularly wrong model. One should note, that this is fully true only for high (infinite) number of output bins, where the discrete output shape can be considered as continuous and unfolding gets the full freedom in its adaptation. In practice though, the unfolding is limited by the data statistics and is only as good as correct is the model distribution within

[5]In addition, number of degrees of freedom is a small number: four for E_T^γ and five for η^γ.

Figure 8.13: Proton - antiproton collision without (a) and with (b) multiparton interaction. Figure taken from [126]

At HERA, MPI can happen in case of resolved events, where partons from both proton

and photon which do not participate in the hard interaction, are still allowed to interact. Both HERA experiments, ZEUS [127] and H1 [128] reported observation of effects that could be assocated with MPI.

The QCD calculations do not include MPI and need to be corrected for its presence. Its influence was modelled with PYTHIA MC with default MPI settings [129, 130] (MC set IV in table 2.3) as compared to PYTHIA MC generated with MPI being switched off (MC set III). The correction factor f_{theor}^{MPI} then was defined:

$$f_{theor,i}^{MPI} = N_i^{MPI}/N_i^{\sim MPI}, \tag{8.50}$$

where N_i^{MPI} is the number of events in bin i generated with MPI and $N_i^{\sim MPI}$ is the same but without MPI.

In the MC PYTHIA model, final state partons are produced with transverse momenta down to the cut off value p_T^{min}. It was shown, that lowering the p_T^{min} value increases the influence of multiparton interaction [129]. For a determination of the MPI uncertainty, the default value $p_T^{min} = 1.9$ GeV was changed between 2.2 and 1.6 GeV. The correction factor, as determined for all the final cross sections, together with its uncertainty is presented in figure 8.14. As expected, direct events accumulated around $x_\gamma^{LO} \sim 1$ stay unaffected by the MPI effect. At the same time it decreases the cross section for the resolved events by even 15% for the lowest x_γ^{LO} bin. Consistently, cross sections in the resolved enhanced phase space are strongly affected, in some parts even correction of 20% must be introduced. The size of the correction increases for the high η^γ and η^{jet} values, dominated by the resolved events. The additional energy flow introduced by MPI affects low energetic jets to a higher degree compared to more energetic ones, the effect which is reflected in relatively higher MPI corrections for low E_T^{jet} and low $p_\perp$.

8.7.2 Isolation definition correction

Both calculations were performed in the phase space that differs from the one studied in this thesis in terms of photon isolation. In case of QCD calculations, the phase space is defined with the cone isolation criteria, where the total transverse energy within a cone of radius one in (η, ϕ) plane[6] surrounding the prompt photon, $E_T^{(\eta,\phi)=1}$, excluding its own energy, is required to be below 10% of the photon energy:

$$\frac{E_T^{(\eta,\phi)=1}}{E_T^\gamma} < 0.1. \tag{8.51}$$

The transition into the phase space with z-based photon isolation (see section 1.5) was performed using correction factor f_{theor}^{isol} determined for each bin i:

$$f_{theor,i}^{isol} = N_i^z/N_i^{cone}, \tag{8.52}$$

where N_i^z is the number of MC events in bin i selected with the z-based isolation criteria, while N_i^{cone} denotes the number of events fulfilling the cone isolation criteria. Figure 8.15

[6]For the (η, ϕ) distance calculation, azimuthal angle is expressed in radians.

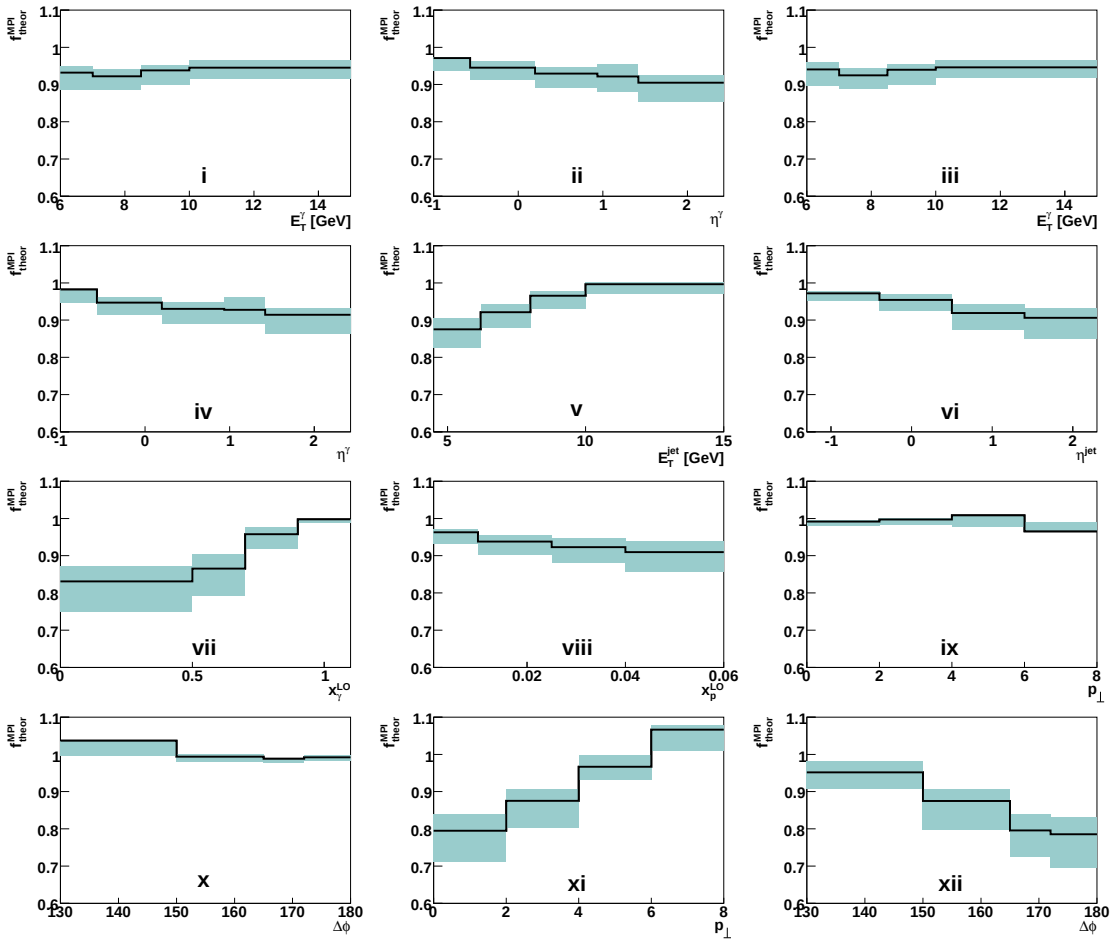

Figure 8.14: Multi parton interaction correction f^{MPI}_{theor}, together with its uncertainty as determined for all the final cross sections. The cross sections codes as defined in table 8.1 are given in relevant places.

presents the f^{isol}_{theor} corrections for all the final cross sections. The correction is in general positive and usually below 5%. It rises to values above 5% in the forward region, which may be associated with the increase of the local particle density. The comparison of corrections in the inclusive and exclusive phase spaces leads to the conclusion that the additional selection of the hadronic jet does not affect the isolation of the photon. The strongest correction of almost 10% can be seen in the lowest x_γ^{LO} bin for the predominantly resolved events. For the high x_γ^{LO} values, where direct events accumulate, as well as for the cross sections in the direct enhanced phase space, the correction is below 1%.

8.7.3 Hadronisation correction

The data measurement is presented as a cross section on a hadron level and parton level calculations must be corrected for hadronisation effects. A hadronisation correction factor

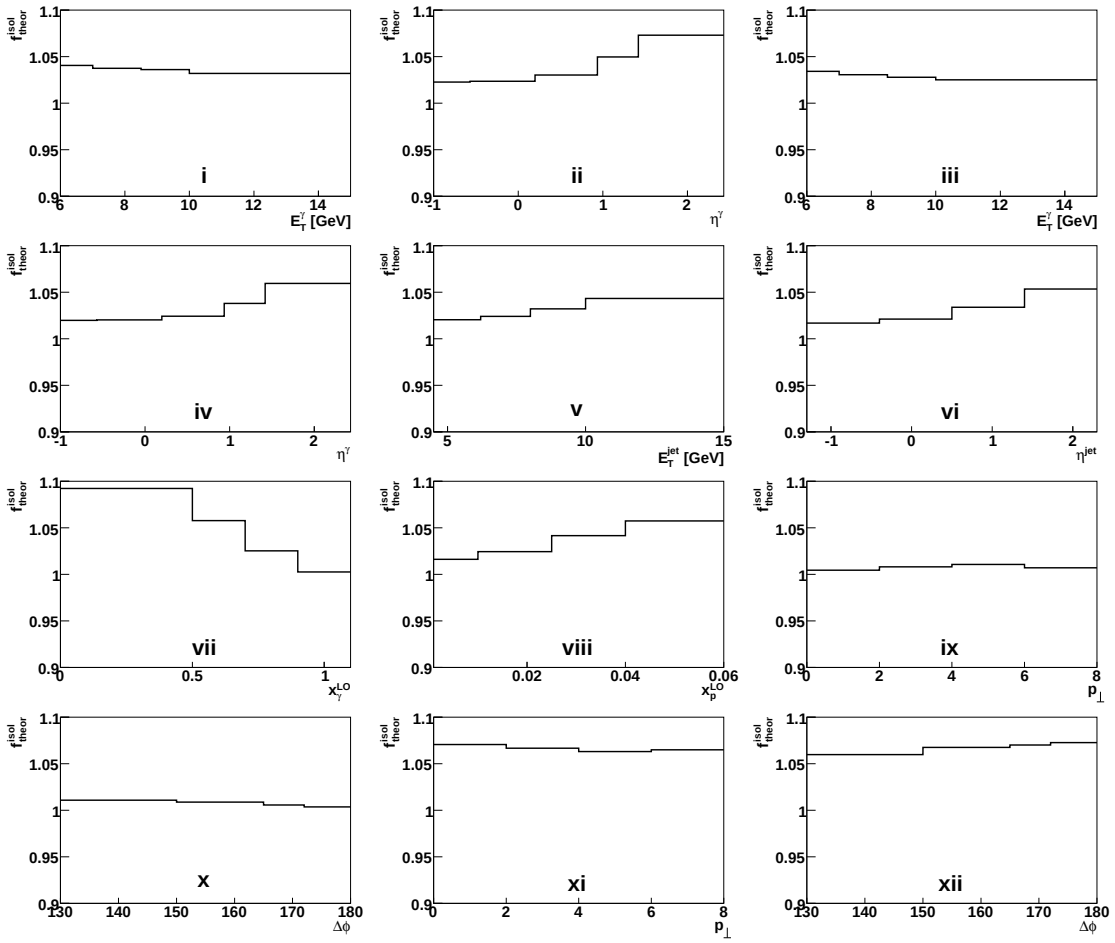

Figure 8.15: Isolation correction f^{isol}_{theor} as determined for all the final cross sections. The cross sections codes as defined in table 8.1 are given in relevant places.

f^{had}_{theor} is determined for each bin i:

$$f^{had}_{theor,i} = N^{had}_i / N^{par}_i, \tag{8.53}$$

where N^{had}_i and N^{par}_i is the number of generated events in the given phase space in bin i on hadron and parton level respectively. Since the hadronisation in MC can be calculated using two competitive methods (see section 2.1.1), the correction factor f^{had}_{theor} is determined with the help of two different Monte Carlo generators. The actual correction factor is taken as the mean value between the correction factor obtained with PYTHIA MC (using Lund fragmentation model) and with HERWIG MC (CDF model), while half of the difference between them determines the corresponding uncertainty.

The correction factors for all the final cross sections is shown in figure 8.16. For the inclusive cross sections, the hadronisation correction is most significant for low transverse energies and in the forward direction. At low energies hadron level jets are statistically wider which may affect the resolution in the isolation of the photon. In the forward region of the detector, the proton remnant creates higher local particle density and thus more sensitivity to the isolation cut. For the exclusive cross section, correction factors obtained

with PYTHIA and HERWIG significantly differ and thus the fragmentation model uncertainty becomes significant, even up to 5% order in the lowest E_T^γ bin. The fragmentation model used in HERWIG systematically lead to stronger hadronisation corrections.

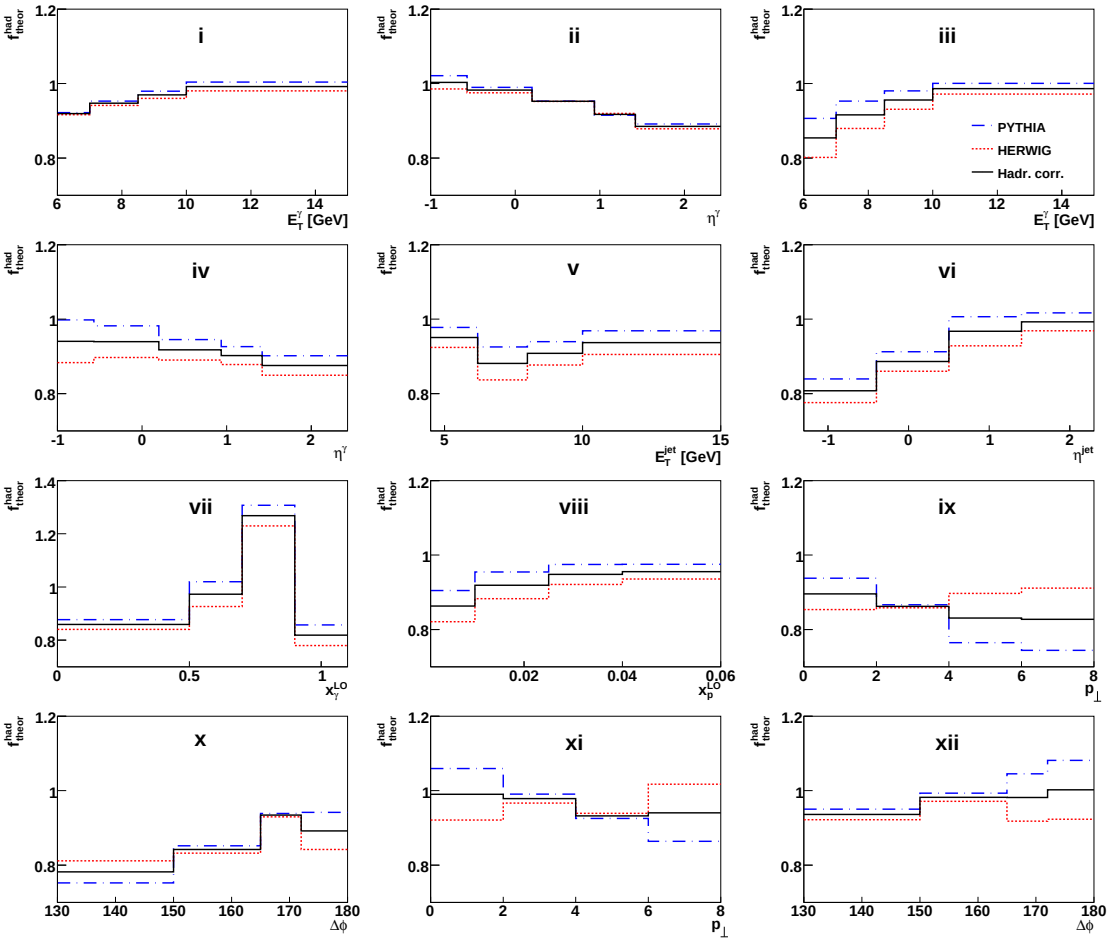

Figure 8.16: Hadronisation correction for all the final cross sections. The dashed-dotted line indicate hadronisation correction obtained with PYTHIA MC, dotted with HERWIG MC and the solid line represents hadronisation correction applied to the calculations. The cross sections codes as defined in table 8.1 are given in relevant places.

8.7.4 Parton density function uncertainty

Another theoretical uncertainty arises from the choice of the PDFs used in the calculation. To evaluate the sensitivity to the chosen PDF set, different MCs were produced (MC sets XI-XV). Three proton parton density functions (CTEQ6L [73], CTEQ5L [26] and MRST04 [27]) and two photon PDFs (AFG [50] and GRV [82]) were studied. For each

analysis bin i a $\pm\delta^{PDF}$ is calculated according to:

$$+\delta_i^{PDF} = \max_j\left(\frac{N_i^j - N_i^{ref}}{N_i^{ref}}\right) \times 100\%, \tag{8.54}$$

$$-\delta_i^{PDF} = \max_j\left(\frac{N_i^{ref} - N_i^j}{N_i^{ref}}\right) \times 100\%, \tag{8.55}$$

where N_i^{ref} is the number of events in the bin i produced with the reference PDF set and index j runs over all the PDF sets, including the reference one.

For the uncertainty applied to the FGH calculation (see section 2.2.1), the reference cross section was chosen to be the one produced with CTEQ6L proton PDF and AFG04 photon PDF, as exactly this set was used during calculations. Since the chosen PDF set consistently produces the lowest cross sections in the studied phase space, the resulting error is highly asymmetric.

The ZL calculation (see section 2.2.2) uses the $k_\perp$ unintegrated PDF, which could not be easily applied to the MC. For that purpose, the compromise solution was found, to use the same PDF uncertainty as evaluated for FGH, but apply it in a symmetric way.

Figure 8.17 presents the PDF uncertainty obtained for the final cross sections, as applied to the both sets of calculations. The choice of the PDF arises to the leading uncertainty source in most of the studied phase space being on average slightly below 10% and reaching even 15% in some bins of resolved enhanced phase space.

Due to possible cancellations, special attention was taken to ensure correct multiplication of the different corrections. The f_{theor}^{MPI} correction, being the first correction applied to the calculation was determined studying the MC on parton level and with a cone isolation criteria. The f_{theor}^{isol} correction was determined with MCs including MPI, but still on parton level. And finally, for the determination of the third correction, f_{theor}^{had}, both MCs were generated with MPI and the cross section phase space was already based on a z variable. In this way, the final correction factor f_{theor} may be calculated with a reduced formula[7]:

$$f_{theor} = f_{theor}^{MPI} \times f_{theor}^{isol} \times f_{theor}^{had} \tag{8.56}$$

$$= N_i^{had,z,MPI}/N_i^{par,cone,\sim MPI}, \tag{8.57}$$

where $N_i^{par,cone,\sim MPI}$ is the number of events in bin i on parton level, with cone based isolation and without MPI (the phase space directly calculated by the QCD calculations), while $N_i^{had,z,MPI}$ is the number of events in bin i on hadron level, with z-based isolation and with MPI (the phase space being the result of the measurement).

All the theory corrections f_{theor}^{MPI}, f_{theor}^{isol}, f_{theor}^{had}, together with their uncertainties and the PDF uncertainty δ^{PDF} are listed in the relevant tables in appendix C.

[7] In practice though, the correction is calculated with an extended triple correction formula, due to the f_{theor}^{had} correction determination method, which includes the usage of two different MC generators.

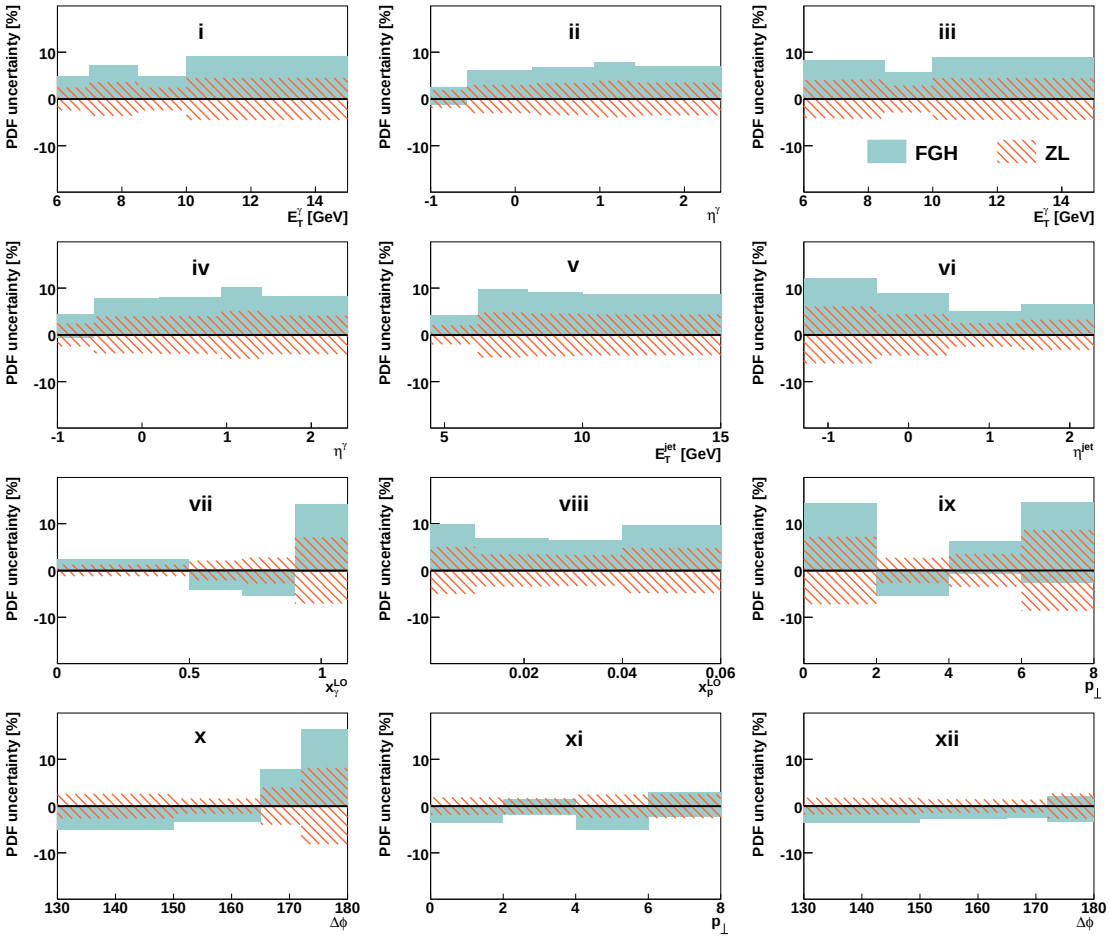

Figure 8.17: Parton density function uncertainty for all the final cross sections. The cross sections codes as defined in table 8.1 are given in relevant places.

Chapter 9

Results

In this chapter results are presented for the production of prompt photons in photoproduction region based on the data taken by the H1 detector with an integrated luminosity of 430 pb^{-1}. Differential cross sections are compared to QCD calculations and to MC predictions introduced previously.

The inclusive prompt photon cross sections are shown and discussed in section 9.1. The differential cross sections in the exclusive photon + jet phase space follow in section 9.2 with special emphasis on the photon - jet correlations. The influence of higher order corrections is studied in section 9.3. Section 9.5 presents the comparison to the previously published results in the accordingly adjusted phase space.

9.1 Inclusive prompt photon production in photoproduction

The measured inclusive prompt photon cross section for the kinematic range specified in table 1.4 is

$$\sigma(ep \to e\gamma X) = 66.9 \pm 1.8\ (\text{stat})\ \pm 8.7\ (\text{syst})\ \text{pb}.$$

Both calculations, corrected for all the effects discussed in section 8.7 predict lower cross sections of $52.1^{+5.3}_{-3.4}$ pb (FGH) and $56.7^{+2.3}_{-3.1}$ pb (LZ), while the MC expectation from PYTHIA is 46.4 pb.

Theoretical uncertainties due to missing higher orders are estimated by simultaneously varying μ_R and μ_F by a factor of 0.5 to 2.0. In addition, the errors on the theoretical predictions include all the uncertainties determined in the section 8.7. All error sources are added in quadrature.

Bin averaged differential cross sections $\mathrm{d}\sigma/\mathrm{d}E_T^\gamma$ and $\mathrm{d}\sigma/\mathrm{d}\eta^\gamma$ for the inclusive prompt photon production are given in tables C-1a and C-2a in appendix C and are presented in the figure 9.1 compared to the prediction from the calculations. Both calculations are below the data, most significantly at low E_T^γ. The LZ calculation gives a reasonable

description of the shape of η^γ, whereas the FGH calculation is significantly below the data for central and backward photons ($\eta^\gamma < 0.9$).

The double differential measurement $\mathrm{d}\sigma^2/\mathrm{d}E_T^\gamma \mathrm{d}\eta^\gamma$ is given in table C-3a and is presented in figure 9.2. In this case, results are also compared to the PYTHIA Monte Carlo prediction. LZ provides a reasonable description of the data with the exception of the lowest E_T^γ bin in the central η^γ ($0.2 < \eta^\gamma < 0.9$) region. The FGH calculation underestimates the cross section in the central ($0.2 < \eta^\gamma < 0.9$) and backward ($\eta^\gamma < -0.6$) region. Here, it is significantly below the data. The prediction from PYTHIA underestimates the measured cross section by roughly 45%, most significantly at low E_T^γ.

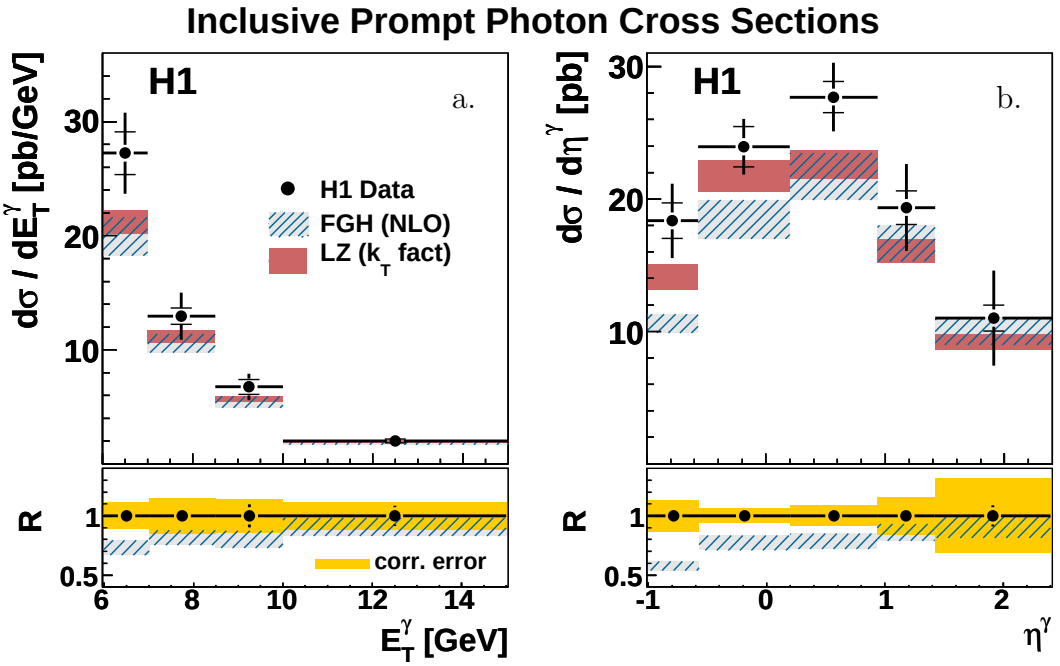

Figure 9.1: The differential inclusive prompt photon cross sections in bins of E_T^γ (a) and η^γ (b). The inner error bars indicate the fully uncorrelated error, while the full error bars contain in addition the correlated error added in quadrature. The measurement is plotted together with the NLO collinear calculation prediction (FGH) and the prediction based on the k_T factorisation approach (LZ). The lower part of the plot presents the ratio R of the FGH prediction to the measurement together with the data points. The error bars associated with the points indicate the fully uncorrelated error, while the yellow band shows the fully correlated error.

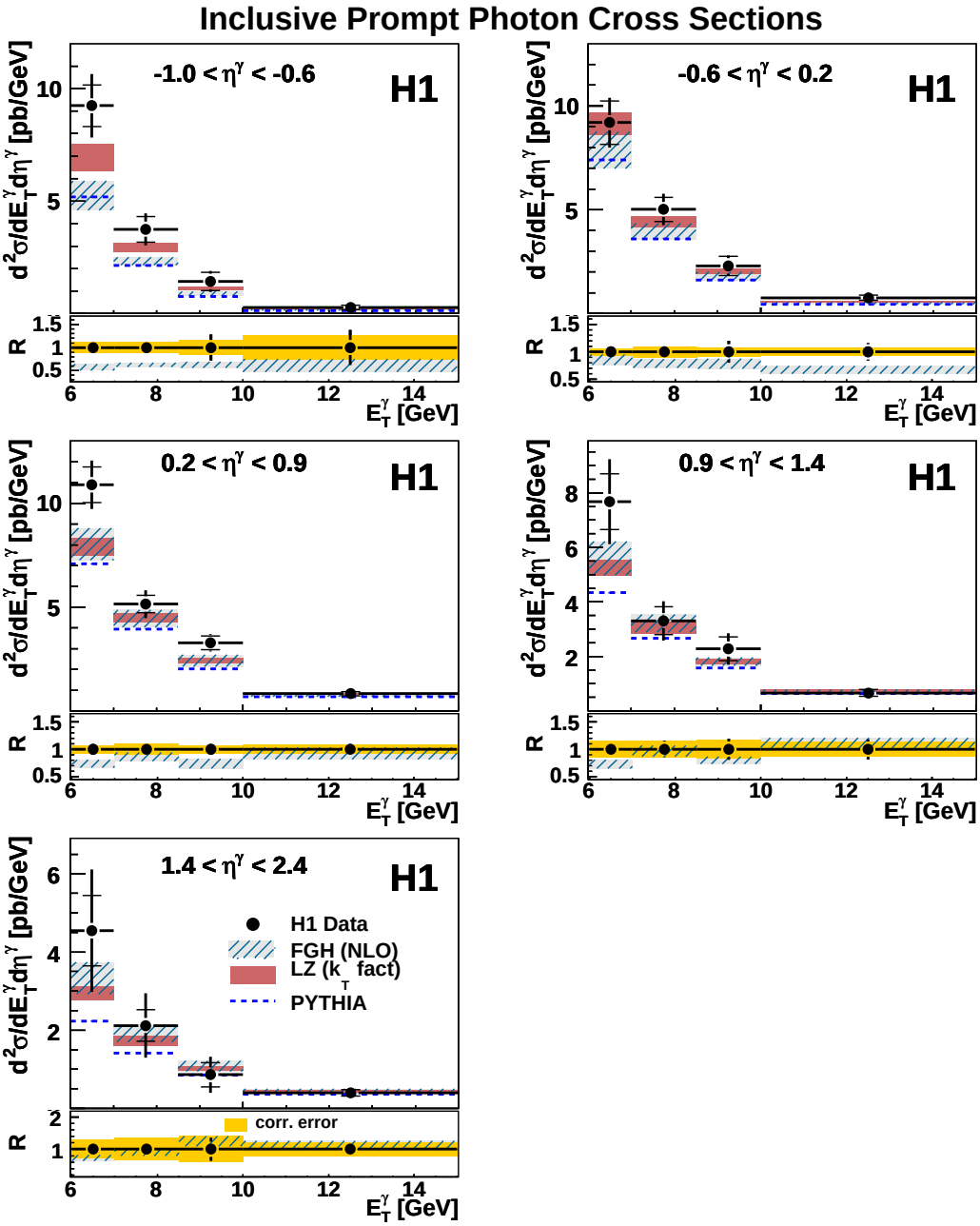

Figure 9.2: The double differential inclusive prompt photon cross sections in bins of E_T^γ and η^γ. The inner error bars indicate the fully uncorrelated error, while the full error bars contain in addition the correlated error added in quadrature. The measurement is plotted together with the NLO collinear calculation prediction (FGH) and the prediction based on the k_T factorisation approach (LZ) and the PYTHIA Mc prediction. The lower part of the plot presents the ratio R of the FGH prediction to the measurement together with the data points. The error bars associated with the points indicate the fully uncorrelated error, while the yellow band shows the fully correlated error.

9.2 Prompt photon + jet production in photoproduction

The cross section for prompt photon production associated with a jet with a given kinematics is studied. The measurement is performed in the *exclusive* phase space, defined in table 1.4. The measured prompt photon plus jet cross section is

$$\sigma(ep \to e\gamma \text{ jet } X) = 50.1 \pm 1.7 \text{ (stat)} \ \pm 6.5 \text{ (syst) pb.}$$

The theoretical calculations predict cross sections of $43.8^{+5.3}_{-1.9}$ pb (FGH) and $49.3^{+4.7}_{-2.1}$ pb (LZ). Both are compatible with the measurement within the errors. The MC expectation of 33.9 pb is again too low.

The differential cross sections in bins of E_T^γ, η^γ, E_T^{jet} and η^{jet} in the exclusive prompt photon + jet phase space are presented in figure 9.3 and tables C-4a, C-5a, C-6a and C-7a. Both calculations give a reasonable description of the E_T^γ and E_T^{jet} cross sections while only the FGH calculation describes η^{jet} of the associated hadronic jet well. Here, the LZ prediction is too high for jets with $\eta^{\text{jet}} < 0.5$. As in the inclusive case, the FGH prediction is too low for $\eta^\gamma < -0.5$.

The presence of a jet allows to use observables which give more insight in the underlying partonic process than in the inclusive case, as defined in section 1.5. Figure 9.4 and tables C-8a and C-9a show the photon plus jet cross section as a function of the estimators x_γ^{LO} and x_p^{LO}. Both distributions are well described by the calculations.

9.2.1 Photon - jet correlations

Figure 9.5 and tables C-10a and C-11a show the cross sections for two observables describing the transverse correlation between the photon and the jet, $p_\perp$ and $\Delta\Phi$. Both variables are particularly sensitive to higher order gluon emission. The sample is split into a sample with $x_\gamma^{LO} > 0.8$ where the direct interaction of a photon with the proton dominates and a sample with $x_\gamma^{LO} < 0.8$ with a significant contribution of events with a resolved photon (see figure 5.7a). Both predictions describe the distributions for $x_\gamma^{LO} > 0.8$ reasonably well but tend to underestimate the tails. For $x_\gamma^{LO} < 0.8$ the $p_\perp$distribution is slightly broader than for $x_\gamma^{LO} > 0.8$, which reflects the increased contributions from events with a resolved photon and from photons radiated from quarks in di-jet events. The FGH calculation poorly describes the $p_\perp$ distribution but gives a reasonable description of the measurement in $\Delta\Phi$ for $x_\gamma^{LO} < 0.8$, except for the highest bin in $\Delta\Phi$. The region $\Delta\Phi \to 180°$ is sensitive to multiple soft gluon radiation which limits the validity of fixed order calculations. The LZ calculation includes multiple soft gluon radiation in the initial state before the hard subprocess and describes $\Delta\Phi > 170°$ but predicts a significantly lower contribution of events in the tails of both distributions.

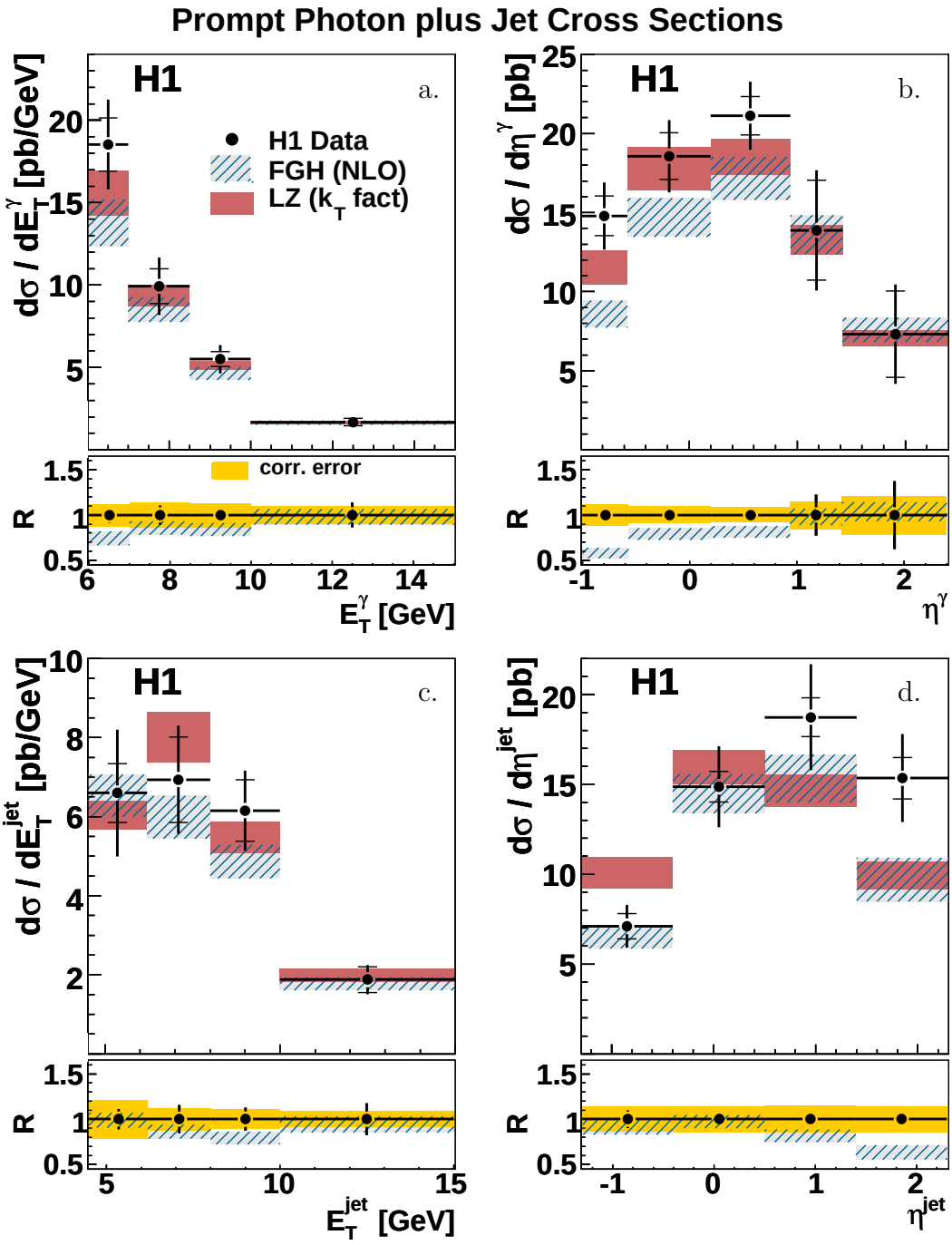

Figure 9.3: The differential exclusive prompt photon + jet cross sections in bins of E_T^γ (a), η^γ (b), E_T^{jet} (c) and η^{jet} (d). The inner error bars indicate the fully uncorrelated error, while the full error bars contain in addition the correlated error added in quadrature. The measurement is plotted together with the NLO collinear calculation prediction (FGH) and the prediction based on the k_T factorisation approach (LZ). The lower part of the plot presents the ratio R of the FGH prediction to the measurement together with the data points. The error bars associated with the points indicate the fully uncorrelated error, while the yellow band shows the fully correlated error.

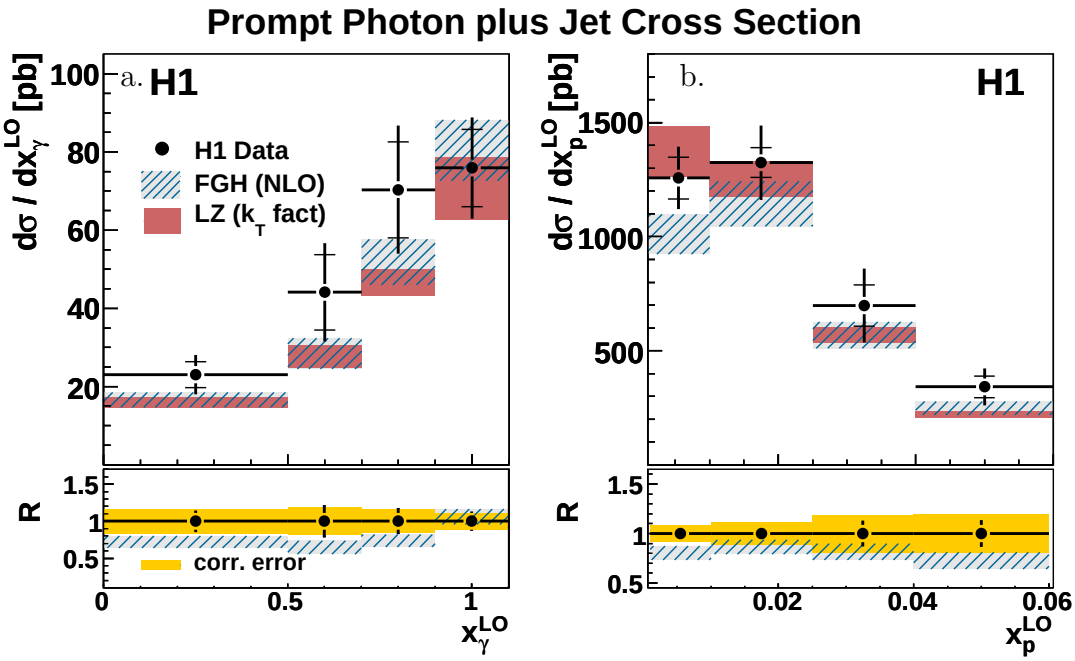

Figure 9.4: The differential exclusive prompt photon + jet cross sections in bins of x_γ^{LO} (a) and x_p^{LO} (b). The inner error bars indicate the fully uncorrelated error, while the full error bars contain in addition the correlated error added in quadrature. The measurement is plotted together with the NLO collinear calculation prediction (FGH) and the prediction based on the k_T factorisation approach (LZ). The lower part of the plot presents the ratio R of the FGH prediction to the measurement together with the data points. The error bars associated with the points indicate the fully uncorrelated error, while the yellow band shows the fully correlated error.

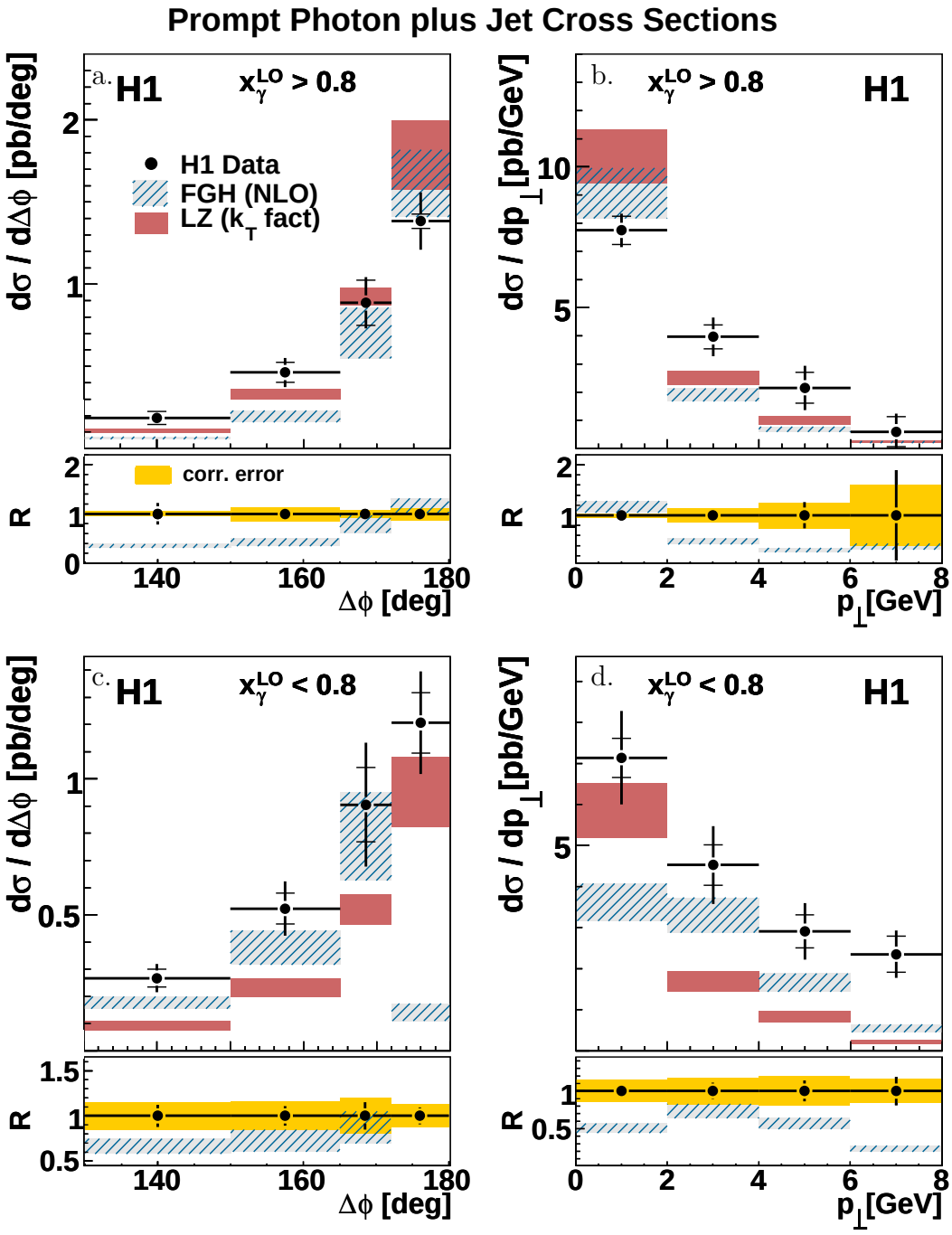

Figure 9.5: The differential prompt photon + jet cross sections in bins of $\Delta\Phi$ (a+c) and $p_\perp$ (b+d) in the direct enhanced phase space (a+b) and in the resolved enhanced phase space (c+d). The inner error bars indicate the fully uncorrelated error, while the full error bars contain in addition the correlated error added in quadrature. The measurement is plotted together with the NLO collinear calculation prediction (FGH) and the prediction based on the k_T factorisation approach (LZ). The lower part of the plot presents the ratio R of the FGH prediction to the measurement together with the data points. The error bars associated with the points indicate the fully uncorrelated error, while the yellow band shows the fully correlated error.

9.3 NLO corrections

The leading order Feymann diagrams for direct and resolved prompt photon production is of order $(\alpha^2\alpha_s^0)$. The FGH calculation (see 2.2.1) is performed to the order $(\alpha^2\alpha_s)$, which includes already the next-to-leading order correction. The correction increases the calculation by roughly 25% in the inclusive prompt photon phase space and by 10% in the exclusive photon + jet phase space. More differential information is presented in figure 9.6 where few chosen measurements are presented together with the LO calculation, the NLO calculation and the calculation without the box diagram (see 1.13a) considered.

The cross section in the inclusive phase space (figures 9.6a and 9.6b) are affected by the correction most significantly at low transverse energies (correction of 32%) and for photons in the forward direction (43%). For the cross sections in the exclusive phase space a very strong positive correction (up to 70%) can be seen for events with the hadronic jet going into the forward direction (figure 9.6d). At the same time, NLO slightly decreases cross sections in the low η^{jet} range. The NLO influence is strongest for the cross sections in bins of x_γ^{LO} (figure 9.6e), where additional gluon radiations shift direct events towards lower x_γ^{LO} values. The correction significantly decreases the cross section in the highest x_γ^{LO} bin (by 40%) but increases it in the other bins.

The box diagram increases calculated cross sections on average by 11%. Its influence is the strongest for low pseudorapidities of the photon, where it increases the cross sections by almost 20%. Low values of x_p^{LO} are also strongly affected (difference of 20%). An interesting feature can be observed for the cross section in bins of x_γ^{LO} variable. Even though the box diagram can not be unambiguously assigned to the direct event class, it contributes only at $x_\gamma^{LO} \approx 1$. Since in the final state there are only two particles leaving the hard interaction, the x_γ^{LO} definition is properly balanced in the same way as it is the case for leading order direct events.

Figure 9.7 presents the same comparison for the cross sections in bins of photon - jet correlation variables in the direct and resolved enhanced phase space separately. In leading order, only back-to-back configuration exists. The photoproduction criteria $Q^2 < 1\,\text{GeV}^2$ together with small intrinsic $\langle k_T \rangle$ of partons in the proton do not yield any sizeable momentum imbalance which could be seen outside of the lowest $p_\perp$ bin of 2 GeV width. The next order correction smears out both distributions. The box diagram contributes mainly in the back-to-back configuration in the direct enhanced phase space.

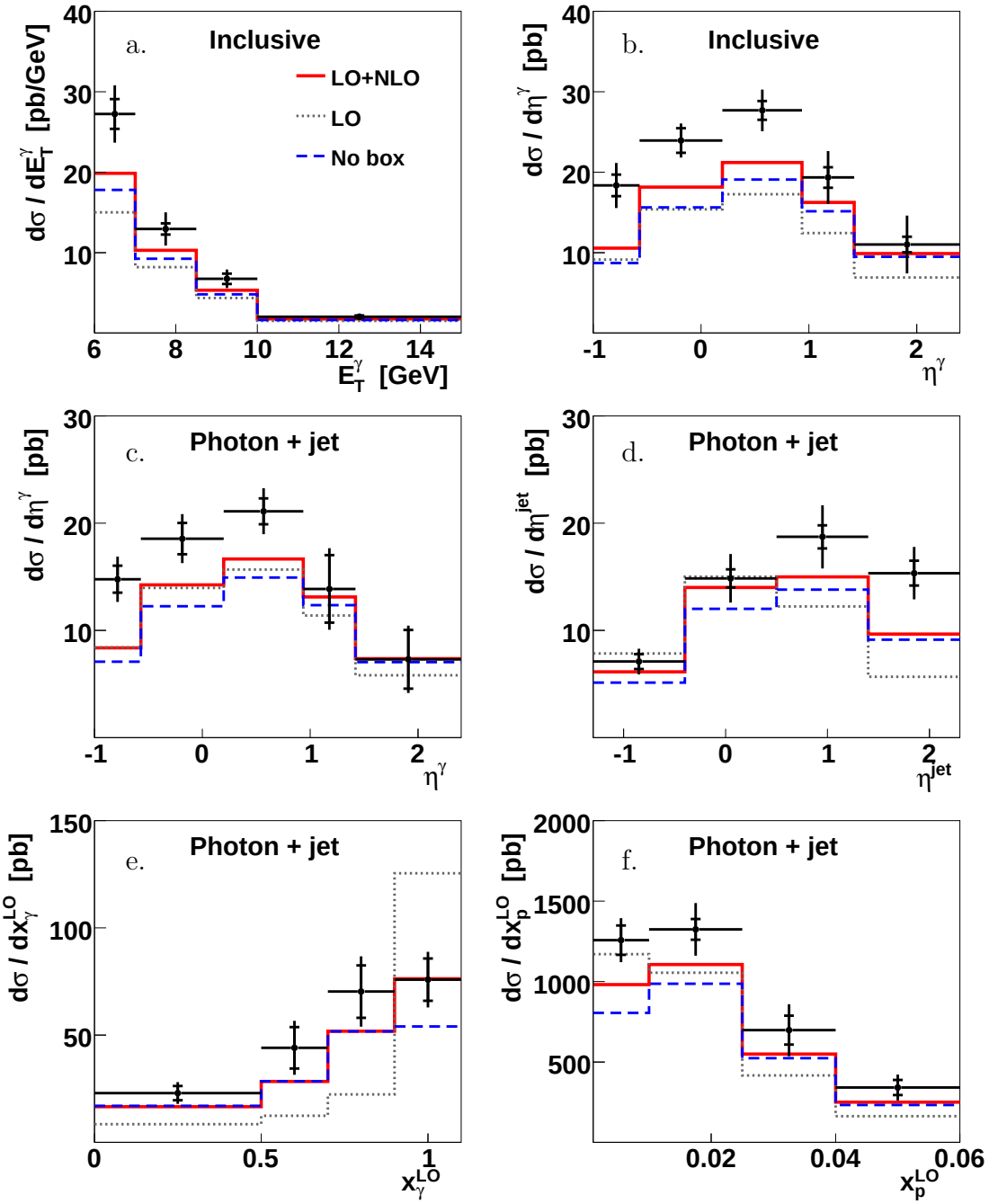

Figure 9.6: The measured cross sections in bins of E_T^γ (a) and η^γ (b) in the inclusive phase space as well as η^γ (c), η^{jet} (d), x_γ^{LO} (e) and x_p^{LO} (f) in the exclusive phase space plotted together with the FGH calculation (LO+NLO - solid red line). The calculation taking into account only leading order diagrams (LO) is plotted as dotted gray histogram and the one taking into account all except the box diagram is plotted as dashed blue histogram.

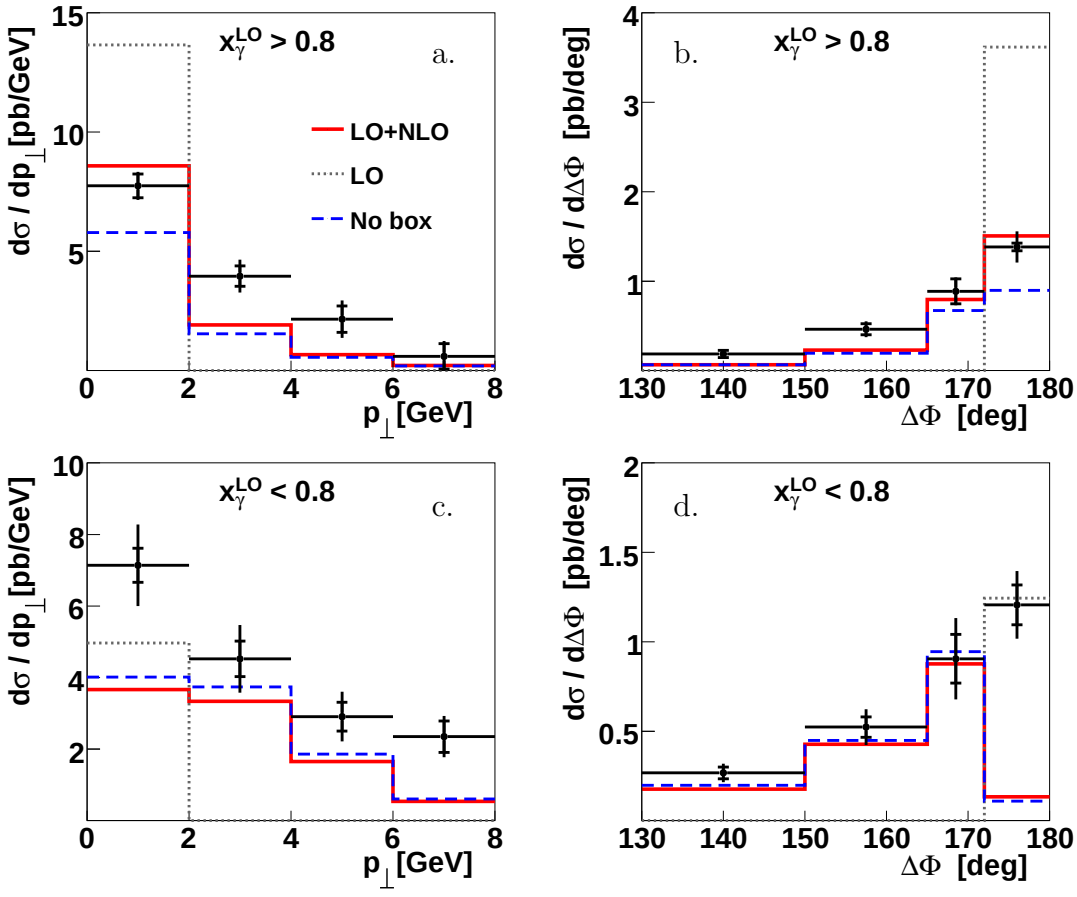

Figure 9.7: The measured cross sections in bins of $p_\perp$ (a, c) and $\Delta\Phi$ (b, d) in the direct (a, b) and resolved (c, d) enhanced phase space plotted together with the FGH calculation (LO+NLO - solid red line). The calculation taking into account only leading order diagrams (LO) is plotted as dotted gray histogram and the one taking into account all except the box diagram is plotted as dashed blue histogram.

9.4 Intrinsic $\langle k_T \rangle$ study

As explained already in the section 1.5, the interesting study can be performed whether the description of the measurement by theory can be improved by introduction of the intrinsic momentum of partons within a proton $\langle k_T \rangle$. Variables $p_\perp$ and $\Delta\Phi$ are chosen to enhance the measurement sensitivity to the size of $\langle k_T \rangle$. Both variables provide the direct information of the initial transverse momentum of the parton entering the hard interaction only in direct photon interaction and only in leading order. For that reason, both distributions are measured in the direct enhanced phase space with the cut $x_\gamma^{LO} > 0.8$ applied. The result of this study is presented in section 9.2.1.

In the PYTHIA MC model the intrinsic $\langle k_T \rangle$ value is a parameter that can be adjusted. Special MC study is performed in order to investigate the possible improvement of the photon - jet correlation description by the theory purely by adjusting the value of this parameter. The measurement was compared to the PYTHIA MC prediction with the $\langle k_T \rangle$ values equal to $0.0, 0.5, 1.0, 1.5$ and $2.0\,\mathrm{GeV}$. The MC has been normalised to the highest bin in $\Delta\Phi$. The cross section in bins of both sensitive variables in the direct enhanced phase space are presented in figure 9.8.

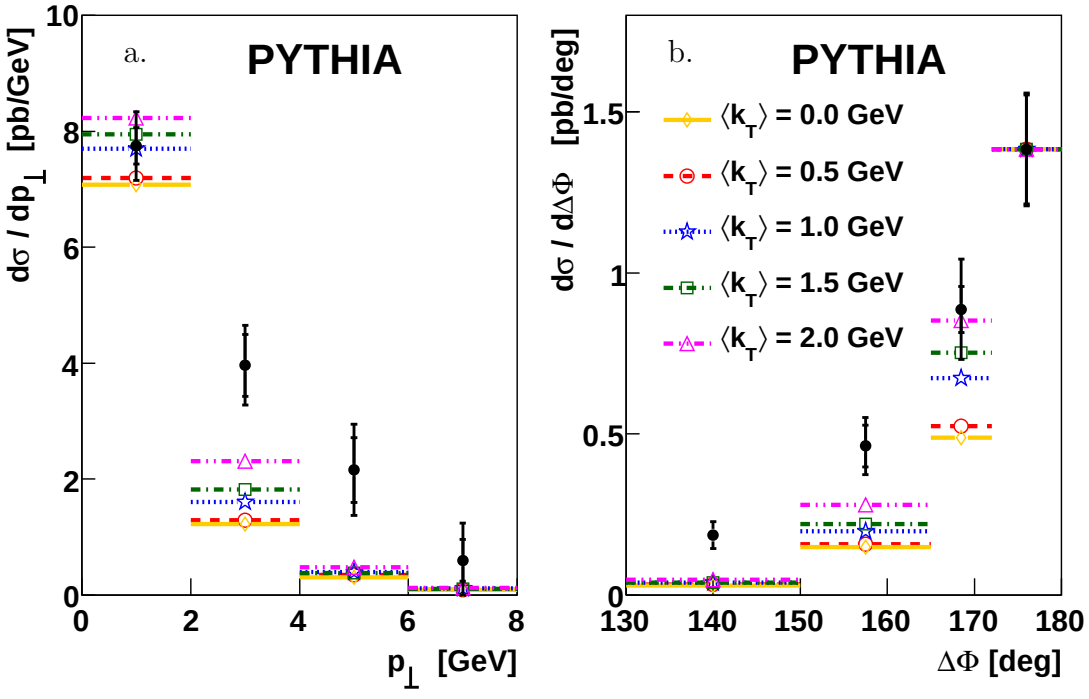

Figure 9.8: The cross section in bins of $p_\perp$ (a) and $\Delta\Phi$ (b) in the direct enhanced phase space compared to the PYTHIA MC prediction produced with variable $\langle k_T \rangle$ value. The MC is normalised to the highest bin in $\Delta\Phi$.

Low value of $\langle k_T \rangle$ leads to high correlation between photon and the jet and just few events in the tails of the distribution of both variables originating mainly from initial and final state gluon emissions. The increase of the intrinsic momentum value introduces intrinsic

transverse momentum imbalance already in leading order and migrates events to the tails of both distributions. Even with $\langle k_T \rangle = 2.0\,\text{GeV}$ though, PYTHIA MC fails to properly describe the tails which may lead to the conclusion that the observed discrepancy can not solely be explained by the missing transverse momentum of the struck parton.

In order to further investigate the observed discrepancy, the measurement is compared to the FGH NLO calculation, additionally corrected for the $\langle k_T \rangle$ effect. For every studied $\langle k_T \rangle$, for every bin i in $p_\perp$ and $\Delta\Phi$ distributions, the correction factor is defined as

$$f_{theor,i}^{\langle k_T \rangle = k} = N_i^{\langle k_T \rangle = k} / N_i^{\langle k_T \rangle = 0}, \tag{9.1}$$

where $N_i^{\langle k_T \rangle = k}$ is the number of MC events in bin i generated with the mean intrinsic transverse momentum $\langle k_T \rangle = k\,\text{GeV}$. The $f_{theor,i}^{\langle k_T \rangle = k}$ factors are used to correct the FGH calculation. This corrected FGH calculation is compared to the measurement in figure 9.9. One could note the improvement in the measurement description with the additional intrinsic $\langle k_T \rangle$, in the $\Delta\Phi$ shape, though the last bin remains underestimated even with $\langle k_T \rangle = 2.0\,\text{GeV}$. A similar conclusion can be made for the distributions in bins of $p_\perp$. Both comparisons indicate that though the contribution in the tails is increased, the intrinsic $\langle k_T \rangle$ can not solely repair the observed discrepancy.

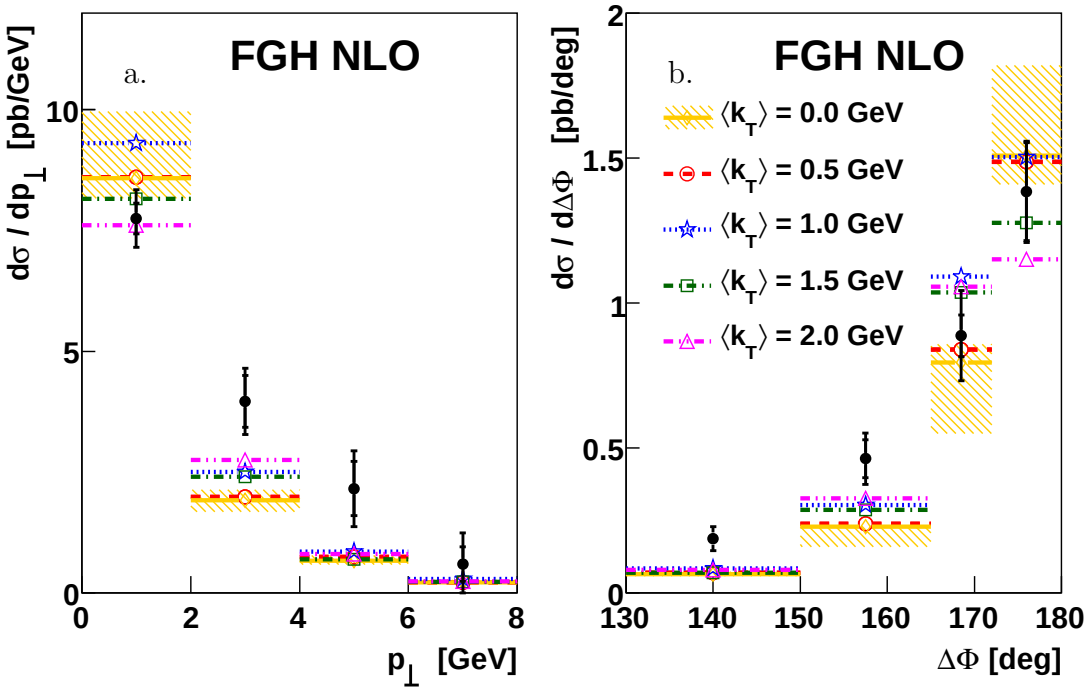

Figure 9.9: The cross section in bins of $p_\perp$ (a) and $\Delta\Phi$ (b) in the direct enhanced phase space compared to the FGH NLO calculation with different corrections for the intrinsic transverse momentum effects. The shaded area indicate the error associated to the original calculation.

In the case of the resolved events, due to higher number of final state particles, the photon - jet correlations are less visible in the resolved enhanced phase space. Roughly

20% of the events in the direct enhanced phase space are actually resolved or radiated events migrating into the high x_γ^{LO} range and contributing significantly to the tails of both distributions. One possible explanation of the observed discrepancy is actually the deficit in modelling of the resolved part of the interaction.

9.5 Comparison to previous measurements

The cross sections were calculated in the phase space of the previous measurements allowing a direct comparison to the published results based on HERA data.

The prompt photon measurement in photoproduction published by the H1 collaboration [35] was based on HERA I data with an integrated luminosity of 105 pb^{-1}. The isolation of the photon is ensured by the cone isolation criteria described in section 8.7.2. The ZEUS collaboration reported the prompt photon + jet cross section measurement with 77 pb^{-1} of HERA I data. Both measurement phase spaces are summarized in the table 9.1. The H1 results were obtained using a signal-background discriminator based on three shower shape variables (R_T, FLF and HCF - see section 6.1), while results of the ZEUS collaboration are based on the signal in the Zeus pre-shower detector [131]. The results of both publications were determined using a fitting procedure described in section 8.6.1.

H1 published phase space	**ZEUS published phase space**
$Q^2 < 1\,\mathrm{GeV}^2$	$Q^2 < 1\,\mathrm{GeV}^2$
$0.2 < y < 0.7$	$0.2 < y < 0.8$
$E_T^{(\eta,\phi)=1}/E_T^\gamma < 0.1$	$z > 0.9$
$5 < E_T^\gamma < 10\,\mathrm{GeV}$	$7 < E_T^\gamma < 16\,\mathrm{GeV}$
$-1.0 < \eta^\gamma < 0.9$	$-0.74 < \eta^\gamma < 1.1$
	$6 < E_T^{\mathrm{jet}} < 17\,\mathrm{GeV}$
	$-1.6 < \eta^{\mathrm{jet}} < 2.4$

Table 9.1: The prompt photon phase space of the H1 [35] and the ZEUS [38] measurements.

For a direct comparison, the phase space of this analysis is adjusted to match the published ones with the only exception of the photon transverse energy range for the H1 published phase space and the photon pseudorapidity range for the ZEUS published phase space. In both cases the difference is preserved in order to visualise the range of the phase space extension. The differential cross sections in bins of E_T^γ in the H1 published phase space and in bins of η^γ in the ZEUS published phase space are presented in figure 9.10.

Figure 9.10a shows the comparison between the published results based on the HERA I data and the HERA II results of this analysis in the properly adjusted phase space. As explained in section 4.3, low transverse energies of the photon are not accessible using H1 HERA II data due to the strong prescaling on the trigger level. Instead, the phase

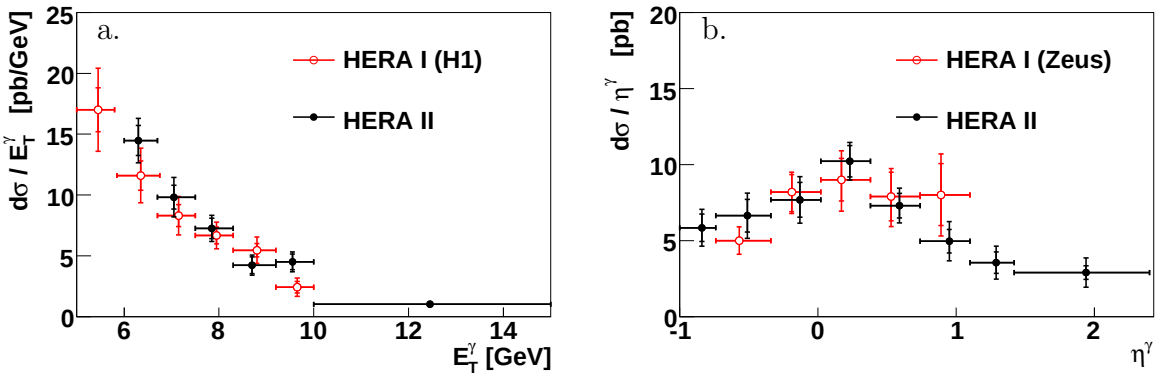

Figure 9.10: The differential cross sections in bins of E_T^γ in the H1 published phase space (a) and in bins of η^γ in the ZEUS published phase space. Both results are compared to the published ones (drawn with open points).

space has been extended towards higher energies. Both results are compatible within errors, taken into account the statistical independence of the used data samples and the difference in the MC simulation (especially in the details of the shower simulation) used in both methods. Still, the new measurement seems to exceed the old one in the lowest transverse energy bin, which is explained by the improved method of trigger energy estimation[1]. The data sample used in this analysis exceeds by a factor of four the statistics on which the published H1 results were based. This is reflected in the small improvement of the measurement precision, though the more advanced unfolding method applied here is known to produces higher error estimates (see section 8.6.2).

The results of the analysis in the Zeus published phase space, shown in figure 9.10b is consistent with the Zeus measurement within the errors. The extension of the accessed pseudorapidity range can be seen, especially into the forward direction. In the nominal phase space of this analysis, the improvement is even more important, as the inelasticity range of the Zeus published phase space removes most of the forward events. The forward photons have never been studied at HERA before.

[1]Due to lack of sufficient DIS data in HERA I running period, monitoring sample has been selected using s71 subtrigger, which is not fully independent of the LAr calorimeter. This leads to an overestimation of a determined trigger efficiency and underestimation of the final cross sections. The effect is strongest in the low transverse energy region, where trigger efficiency is of the highest importance. Also the L4 trigger condition, affecting the low energies has not been fully taken into account in the previous H1 measurement.

Chapter 10

Conclusions and outlook

The prompt photons originating from the hard interaction are a powerful test of QCD. Compared to measurement relying on jets, they are believed to be less sensitive to the details of hadronisation models which introduce large theoretical uncertainties. The more difficult background situation however, which demands complex signal extraction methods leads to an increase of the systematic error. Also statistical error is higher since the photon cross section is suppressed by the order α as compared to the jet cross sections. The prompt photon analysis, though challenging and intrinsically burdened with higher uncertainty, provides a complementary measurement to results relying on jets.

The cross sections for prompt photon production in photoproduction is measured using data taken by the H1 detector with the full HERA II integrated luminosity of 340 pb^{-1}. The study provides the first ever measurement of photons in such a forward direction at HERA. Compared to the previous measurement performed by the H1 collaboration [35], the phase space has been also extendend into lower event inelasticities. Prompt photons with the transverse energy $6.0 < E_T^\gamma < 15.0\,\mathrm{GeV}$, pseudorapidity $-1.0 < \eta^\gamma < 2.4$ and isolation $z > 0.9$ in photoproduction region $Q^2 < 1\,\mathrm{GeV}^2$ and event inelasticity region $0.1 < y < 0.7$ have been studied in the photon inclusive phase space and together with the accompanying hadronic jet with transverse energy $E_T^{\mathrm{jet}} > 4.5\,\mathrm{GeV}$ and pseudorapidity $-1.3 < \eta^{\mathrm{jet}} < 2.3$. The complex unfolding method has been developed in order to perform the measurement insensitive to possible model inaccuracies. The cross sections are compared to QCD calculations based on collinear factorisation in NLO [39, 75] and based on the k_T factorisation approach [80].

In the inclusive prompt photon phase space the $\mathrm{d}\sigma/\mathrm{d}E_T^\gamma$, $\mathrm{d}\sigma/\mathrm{d}\eta^\gamma$ as well as double differential $\mathrm{d}\sigma^2/\mathrm{d}E_T^\gamma\mathrm{d}\eta^\gamma$ cross sections are found to be underestimated by both predictions most significantly for low transverse energies of the photon. Furthermore, FGH calculation fails to describe the shape of the η^γ, being significantly too low for $\eta^\gamma < 0.9$. The cross sections predicted by the PYTHIA Monte Carlo generator is underestimates the measurement by 40% on average, most significantly at low transverse energies.

The differential cross sections $\mathrm{d}\sigma/\mathrm{d}E_T^\gamma$, $\mathrm{d}\sigma/\mathrm{d}\eta^\gamma$, $\mathrm{d}\sigma/\mathrm{d}E_T^{\mathrm{jet}}$, $\mathrm{d}\sigma/\mathrm{d}\eta^{\mathrm{jet}}$ are studied in the photon + jet exclusive phase space. Both calculations describe the rate of prompt photons in the exclusive case somehow better than in the inclusive phase space. This observation

is consistent with the one made in the previous H1 publication [35]. The improvement can be associated with the lower NLO corrections in the exclusive case, which may suggest that the higher order corrections for inclusive prompt photon cross sections may still be significant.

The cross sections in bins of x_γ^{LO} and x_p^{LO} variables were also measured and found to be consistent with the predictions within errors.

Correlation between the prompt photon and the hadronic jet were studied in the part of the phase space with enhanced contribution from direct events ($x_\gamma^{LO} > 0.8$) and the complementary one ($x_\gamma^{LO} < 0.8$). In both cases, the predicted rate of events with the photon not being fully balanced by the leading hadronic jet is below the measured cross sections. Higher order corrections are expected to introduce additional imbalance, possibly improving the calculations. The fraction of events in the tails of the studied distributions may also be increased by introduction of the additional intrinsic $\langle k_T \rangle$ of the partons entering the hard interaction. The comparison of the measured cross sections in the direct enhanced phase space with the QCD calculations corrected for the $\langle k_T \rangle$ effects leads to the conclusion that though the improvement is clearly visible, the intrinsic $\langle k_T \rangle$ can not solely remove the observed discrepancy.

The difference in the azimuthal angle between prompt photon and the accompanying jet $\Delta\Phi$ in the exclusive phase space with additional requirement $x_\gamma^{LO} < 0.8$ shows vulnerability of the FGH calculation. The back-to-back region, $\Delta\Phi \to 180°$, is sensitive to multiple soft gluon radiation which limits the validity of fixed order calculations. The LZ calculation includes multiple gluon radiation in the initial state before the hard subprocess and describes problematic region significantly better.

The prompt photon measurement presented in this thesis may still be improved. Most significantly, higher statistics of monitor sample which consists of pure photons is needed to improve the cluster shape description and decrease the leading systematic error. The limited statistics of deeply virtual Compton scattering events in the kinematic range of this analysis demands finding more suitable source of the photon monitor sample. The errors of this measurement are low enough to resolve the deficits in the QCD calculations applied for the comparison.

The conclusions drawn from this study may become stronger by improving the theoretical understanding of the prompt photon production. Especially, better understanding of the hadronisation model would help in decreasing the error of the predictions. The new parton density functions coming from combined H1 and ZEUS data will help with calculation of more precise predictions. The sizeable NLO corrections lead to the conclusion that higher orders should be investigated in future. The box diagram, significantly contributing to the prompt photon cross sections, on the level of 10%, should be taken into account in all the calculations, as well as contribution from the quark-to-photon fragmentation function. The latter should also be studied in more details.

Appendices

A Analysis Binning

Bin	Label	η range	E_T range [GeV]	Bin	Label	η range	E_T range [GeV]
1	CB1-E_1	[-1.0; -0.6]	[5.0; 5.5]	35	FB1-E_1	[0.95; 1.44]	[5.0; 5.5]
2	CB1-E_2	[-1.0; -0.6]	[5.5; 6.0]	36	FB1-E_2	[0.95; 1.44]	[5.5; 6.0]
3	CB1-E_3	[-1.0; -0.6]	[6.0; 6.7]	37	FB1-E_3	[0.95; 1.44]	[6.0; 6.7]
4	CB1-E_4	[-1.0; -0.6]	[6.7; 7.5]	38	FB1-E_4	[0.95; 1.44]	[6.7; 7.5]
5	CB1-E_5	[-1.0; -0.6]	[7.5; 8.5]	39	FB1-E_5	[0.95; 1.44]	[7.5; 8.5]
6	CB1-E_6	[-1.0; -0.6]	[8.5; 10.0]	40	FB1-E_6	[0.95; 1.44]	[8.5; 10.0]
7	CB1-E_7	[-1.0; -0.6]	[10.0; 15.0]	41	FB1-E_7	[0.95; 1.44]	[10.0; 15.0]
8	CB1-E_8	[-1.0; -0.6]	[15.0; 20.0]	42	FB1-E_8	[0.95; 1.44]	[15.0; 20.0]
19	CB2-E_1	[-0.6; 0.2]	[5.0; 5.5]	43	FB2-E_1	[1.44; 1.76]	[5.0; 5.5]
20	CB2-E_2	[-0.6; 0.2]	[5.5; 6.0]	44	FB2-E_2	[1.44; 1.76]	[5.5; 6.0]
21	CB2-E_3	[-0.6; 0.2]	[6.0; 6.7]	45	FB2-E_3	[1.44; 1.76]	[6.0; 6.7]
22	CB2-E_4	[-0.6; 0.2]	[6.7; 7.5]	46	FB2-E_4	[1.44; 1.76]	[6.7; 7.5]
23	CB2-E_5	[-0.6; 0.2]	[7.5; 8.5]	47	FB2-E_5	[1.44; 1.76]	[7.5; 8.5]
24	CB2-E_6	[-0.6; 0.2]	[8.5; 10.0]	48	FB2-E_6	[1.44; 1.76]	[8.5; 10.0]
25	CB2-E_7	[-0.6; 0.2]	[10.0; 15.0]	49	FB2-E_7	[1.44; 1.76]	[10.0; 15.0]
26	CB2-E_8	[-0.6; 0.2]	[15.0; 20.0]	50	FB2-E_8	[1.44; 1.76]	[15.0; 20.0]
27	CB3-E_1	[0.2; 0.95]	[5.0; 5.5]	51	IF1-E_1	[1.76; 2.43]	[5.0; 5.5]
28	CB3-E_2	[0.2; 0.95]	[5.5; 6.0]	52	IF1-E_2	[1.76; 2.43]	[5.5; 6.0]
29	CB3-E_3	[0.2; 0.95]	[6.0; 6.7]	53	IF1-E_3	[1.76; 2.43]	[6.0; 6.7]
30	CB3-E_4	[0.2; 0.95]	[6.7; 7.5]	54	IF1-E_4	[1.76; 2.43]	[6.7; 7.5]
31	CB3-E_5	[0.2; 0.95]	[7.5; 8.5]	55	IF1-E_5	[1.76; 2.43]	[7.5; 8.5]
32	CB3-E_6	[0.2; 0.95]	[8.5; 10.0]	56	IF1-E_6	[1.76; 2.43]	[8.5; 10.0]
33	CB3-E_7	[0.2; 0.95]	[10.0; 15.0]	57	IF1-E_7	[1.76; 2.43]	[10.0; 15.0]
34	CB3-E_8	[0.2; 0.95]	[15.0; 20.0]	58	IF1-E_8	[1.76; 2.43]	[15.0; 20.0]

Table A-1: The binning used for multivariate analysis.

Bin	D_{IN}	$E^{\gamma}_{T,IN10}$[GeV]	η^{γ}_{IN10}	$E^{\gamma}_{T,IN5}$[GeV]	$\eta^{\gamma}_{IN,8}$
1	$[0.0; 0.2]$	$[5.5; 5.75]$	$[-1.1; -0.8]$	$[5.5; 6.0]$	$[-1.1; -0.57]$
2	$[0.2; 0.4]$	$[5.75; 6.0]$	$[-0.8; -0.57]$	$[6.0; 6.9]$	$[-0.57; -0.2]$
3	$[0.4; 0.6]$	$[6.0; 6.2]$	$[-0.57; -0.2]$	$[6.9; 7.3]$	$[-0.2; 0.2]$
4	$[0.6; 0.8]$	$[6.2; 6.5]$	$[-0.2; 0.2]$	$[7.3; 8.2]$	$[0.2; 0.5]$
5	$[0.8; 1.0]$	$[6.5; 6.9]$	$[0.2; 0.5]$	$[8.2; 17.0]$	$[0.5; 0.95]$
6		$[6.9; 7.3]$	$[0.5; 0.95]$		$[0.95; 1.2]$
7		$[7.3; 8.2]$	$[0.95; 1.2]$		$[1.2; 1.5]$
8		$[8.2; 9.5]$	$[1.2; 1.42]$		$[1.5; 2.45]$
9		$[9.5; 11.0]$	$[1.42; 1.76]$		
10		$[11.0; 17.0]$	$[1.76; 2.45]$		

Table A-2: The input binning used for the unfolding

Bin	$E^{jet}_{T,IN}$[GeV]	η^{jet}_{IN}	$x_{\gamma,IN}$	$x_{p,IN}$	$p_{\perp,IN}$[GeV]	$\Delta\phi_{IN}[^{\circ}]$
1	$[4.5; 4.8]$	$[-1.3; -0.8]$	$[0.0; 0.3]$	$[0001; 0.006]$	$[0.0; 1.0]$	$[120; 140]$
2	$[4.8; 5.2]$	$[-0.8; -1.3]$	$[0.3; 0.5]$	$[0.006; 0.01]$	$[1.0; 2.0]$	$[140; 150]$
3	$[5.2; 6.0]$	$[-0.8; -0.4]$	$[0.5; 0.6]$	$[0.01; 0.015]$	$[2.0; 3.0]$	$[150; 160]$
4	$[6.0; 7.0]$	$[-0.4; 0.1]$	$[0.6; 0.7]$	$[0.015; 0.025]$	$[3.0; 4.0]$	$[160; 170]$
5	$[7.0; 8.0]$	$[0.1; 0.6]$	$[0.7; 0.8]$	$[0.025; 0.03]$	$[4.0; 5.0]$	$[170; 175]$
6	$[8.0; 10.0]$	$[0.6; 1.3]$	$[0.8; 0.9]$	$[0.03; 0.04]$	$[5.0; 6.0]$	$[175; 178]$
7	$[10.0; 12.0]$	$[1.3; 1.7]$	$[0.9; 1.0]$	$[0.04; 0.05]$	$[6.0; 7.0]$	$[178; 180]$
8	$[12.0; 17.0]$	$[1.7; 2.0]$	$[1.0; 1.1]$	$[0.05; 0.07]$	$[7.0; 8.0]$	
9		$[2.0; 2.3]$	$[1.1; 1.2]$		$[8.0; 9.0]$	
10			$[1.2; 1.3]$			

Table A-3: The input binning used for the unfolding (cont.)

Bin	$E^{\gamma}_{T,OUT4}$[GeV]	η^{γ}_{OUT5}	$E^{\gamma}_{T,OUT2}$[GeV]
u	$[0.0; 6.0]$	$[-\infty; -1.0]$	$[0.0; 6.0]$
1	$[6.0; 7.0]$	$[-1.0; -0.6]$	$[6.0; 8.5]$
2	$[7.0; 8.5]$	$[-0.6; 0.2]$	$[8.5; 15.0]$
3	$[8.5; 10.0]$	$[0.2; 0.94]$	
4	$[10.0; 15.0]$	$[0.94; 1.42]$	
5		$[1.42; 2.4]$	
o	$[15.0; +\infty]$	$[1.42; +\infty]$	$[15.0; +\infty]$

Table A-4: The binning used for cross section definition and as unfolding output.

Bin	$E^{jet}_{T,OUT}$[GeV]	η^{jet}_{OUT}	$x_{\gamma,OUT}$	$x_{p,OUT}$	$p_{\perp,OUT}$[GeV]	$\Delta\phi_{OUT}[^{\circ}]$
u	$[0.0; 4.5]$	$[-\infty; -1.3]$	$[-\infty; 0.0]$	$[0.0; 0.001]$		$[0; 150]$
1	$[4.5; 6.2]$	$[-1.3; -0.4]$	$[0.0; 0.5]$	$[0.001; 0.01]$	$[0.0; 2.0]$	$[130; 150]$
2	$[6.2; 8.0]$	$[-0.4; 0.5]$	$[0.5; 0.7]$	$[0.01; 0.025]$	$[2.0; 4.0]$	$[150; 165]$
3	$[8.0; 10.0]$	$[0.5; 1.4]$	$[0.7; 0.9]$	$[0.025; 0.04]$	$[4.0; 6.0]$	$[165; 172]$
4	$[10.0; 15.0]$	$[1.4; 2.3]$	$[0.9; 1.1]$	$[0.04; 0.06]$	$[6.0; 8.0]$	$[172; 180]$
o	$[15.0; +\infty]$	$[2.3; +\infty]$	$[1.1; +\infty]$	$[0.06; +\infty]$	$[8.0; +\infty]$	

Table A-5: The binning used for cross section definition and as unfolding output (cont.)

Bin	$E^{\gamma}_{T,SIDE}$ [GeV]	η^{γ}_{SIDE}
1	$[6.0; 9.0]$	$[-1.0; 0.2]$
2	$[9.0; 15.0]$	$[0.2; 1.4]$
3		$[1.4; 2.4]$

Table A-6: The binning used for the definition of Migration and Side Bins.

B Alternative classifier definitions

For the final cross section determination, the signal - background discrimination is executed with the help of the discriminator calculated with the maximum likelihood MVA method. The method itself was introduced in the section 6.3. There are other, usually more advanced methods, to calculate suitable discriminator and some most well known are shortly discussed in this appendix.

- **Range Search**
  The obvious generalisation of the maximum likelihood method explained in section 6.3 is the usage of multidimensional probability density function. Instead of multiplying n_{var} one-dimensional probability density functions, introducing unwanted assumption of statistical independence, one can use one n_{var}-dimensional probability density function. For this purpose huge training sample is needed to sufficiently populate n_{var} dimensional phase space. In the Range Search [132] approach, the probability density function is locally estimated by counting signal (n_S) and background (n_B) events in the volume V within the n_{var} dimensional phase space around the test event i. The probability density estimate is than calculated as
  $$D^{RS}(i) = \frac{1}{1 + \frac{N_S n_B(i,V)}{N_B n_S(i,V)}} \tag{1}$$
  where $N_{S(B)}$ is the total number of signal (background) events in the training sample.

  The range search method naturally deals with even complex, non linear correlations between input variables. The performance can be enhanced by preprocessed usage of the binary search tree [133] to store the training sample. For higher number of input variables though, the high computing time and great demand for a computing memory becomes an important limitation.

- **k-Nearest Neighbour**
  Similar to Range Search method, k-Nearest neighbour [134] approach tries to estimate the multidimensional probability density function. Unlike the previous though, instead of counting events in the fixed volume V, it takes into account k closest events in the training sample. Its relatively best performance is achieved for a problems with irregular boundary separating signal and background samples. The technical similarity to the range search method implies similar drawbacks.

- **Fisher discriminant**
  In the Fisher discriminant method [135], event selection is performed in the transformed variable phase space with exactly zero linear correlations. The analysis then determines the axis in such transformed multidimensional phase space, such that, signal and background distributions projections upon this axis maximises the separation.

  The Fisher discriminant in spite of its simplicity may be a competitive choice. Especially it is a powerful tool for a Gaussian input variables distributions with linear correlations.

- **Neural Network**
  Neural network pattern recognition method uses a collection of interconnected neurons with each neuron producing a certain response for a given set of activation [136]. In general, each neuron has a defined response function with free parameters to adapt. In the easiest case the network consists of n_{var} input neurons all connected to one output neuron. More advanced networks introduce one or more hidden neuron layers between the input and output layers. The Weierstrass theorem [137] claims that a single hidden layer with high enough number of neurons is sufficient to approximate a given continuous response function to any desired precision. More complex networks, with more than one hidden layers, are sometimes able to achieve the same performance with lower number of neurons. Weights and coefficients for each neuron are determined with "back propagation" method [138]. Their iterative adaptation is based on the difference between actual output and the desired one.

  The relative flexibility of the neural network and in general high performance lead to its wide acceptance within the high energy community as well as outside of it. This is true even in spite of methods complexity. Though possible correlations between input variables do not cause efficiency dimunision, the method is quite sensitive to overtraining [116].

- **Support Vector Machine**
  The support vector machine ideology and implementation is quite complex. Here only a brief overview is given while details can be found in [139]. The idea of the method is to build a hyperplane separating signal and background using only a minimal subset of all training events (vectors). The chosen events are called "support vectors" and removal all other events from the training sample and retraining method leads to the same classifier. The nonlinear extension is performed by mapping the n_{var} dimensional phase space onto higher dimensional phase space in which the linear signal background separation is possible. The higher dimensional mapping can be eliminated by usage of nonlinear kernel functions for the building of the boundary hyperspace.

  Correct usage of the vector machine method lead to performance competitive to the neural network or multidimensional probability density estimation. The complexity and computing inefficiency makes it though rather rarely used.

- **Boosted Decision Tree**
  The decision tree method is based on consequently splitting the input phase space into small regions (leafs) with assigned signal or background status. The branching is performed along a single input variable a time that maximise the separation and repeated until the desired purity is achieved. The decision tree built in this way is relatively quick and transparent. The improvement of the performance is achieved by repeating of the whole tree growing algorithm with enhanced weight of misclassified events, the method known as *boosting* [140]. The whole forest of boosted trees is in the end linearly combined. Boosting stabilises the response of the decision trees with respect to statistical fluctuations in the training sample and generally improve the signal and background separation. Relatively high sensitivity

to overtraining may be partially removed by the *pruning* (removing of statistically insignificant nodes) of decision trees.

The advantage of the boosted decision tree method is its transparency and little tuning requirement. Inclusion of poorly discriminating input variables does not hurt the method in any way, while it may cause problems for other methods. Though in theory boosted decision tree should underperform in comparison to other, more theoretically advanced methods, in practice in many cases it is working better because either there are not enough training events to properly populate multi dimensional probability density function or the neural network architecture has not been correctly optimised. The main advantage of boosted decision tree method, in spite of its slowness and memory demand, is the simplicity and easiness of usage.

All the methods described above have been to some extend trained and tested for their possible application in this analysis. All of them appeared to explore comparable discriminating performance, with some being significantly slower. The results obtained with two most efficient methods: maximum likelihood and neural network agree within 2%.

One way to compare the performance of the different classifiers is to plot signal efficiency and background rejection determined for any possible cut introduced on the discriminator. The example of such a comparison for the first MVA bin in E_T^γ and the **CB1** wheel is presented in figure 1. The further the graph tends towards the upper right corner (high signal efficiency, high background rejection power), the better the classifier. In the case of this analysis though all the classifiers perform equally well with the sole exception of the Support Vector Machine method, which probably was not tuned well enough.

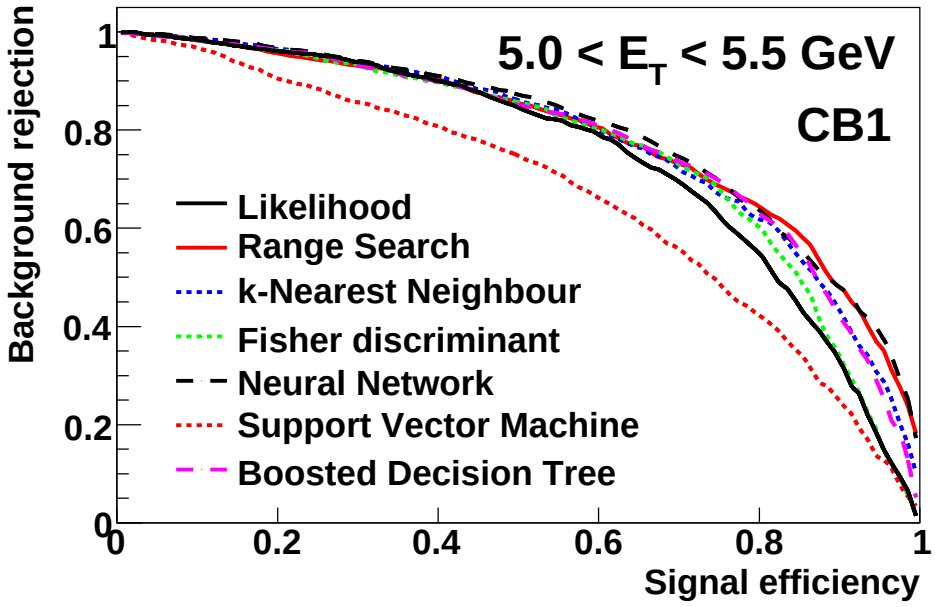

Figure 1: The signal efficiency versus background rejection graph for all studied classifiers in the first MVA E_T bin and the **CB1** wheel.

C Cross Section Tables

η^γ		$\mathrm{d}\sigma/\mathrm{d}\eta^\gamma$	uncorr.	corr.
			[pb]	
$[-1.00$	$-0.57]$	18.35	± 1.34	± 2.47
$[-0.57$	$0.20]$	23.94	± 1.53	± 1.46
$[0.20$	$0.94]$	27.69	± 1.18	± 2.31
$[0.94$	$1.42]$	19.35	± 1.27	± 3.02
$[1.42$	$2.40]$	11.01	± 0.97	± 3.46

Table C-1a: Differential cross sections in bins of η^γ for inclusive prompt photon production in the phase space specified in the table 1.4.

η^γ		f^{MPI}_{theor}	f^{isol}_{theor}	f^{had}_{theor}	δ^{PDF}	
$[-1.00$	$-0.57]$	$0.97^{+0.00}_{-0.03}$	1.02	1.00 ± 0.02	$+2\%$	-1%
$[-0.57$	$0.20]$	$0.95^{+0.02}_{-0.03}$	1.02	0.98 ± 0.01	$+6\%$	-0%
$[0.20$	$0.94]$	$0.93^{+0.02}_{-0.04}$	1.03	0.95 ± 0.00	$+6\%$	-0%
$[0.94$	$1.42]$	$0.92^{+0.03}_{-0.04}$	1.05	0.92 ± 0.00	$+7\%$	-0%
$[1.42$	$2.40]$	$0.90^{+0.02}_{-0.05}$	1.07	0.89 ± 0.01	$+7\%$	-0%

Table C-1b: Corrections applied to the calculations in bins of η^γ.

E_T^γ [GeV]		$\mathrm{d}\sigma/\mathrm{d}E_T^\gamma$ [pb/[GeV]]	uncorr.	corr.
[6.00	7.00]	27.24	±1.86	±3.04
[7.00	8.50]	12.94	±0.71	±1.94
[8.50	10.00]	6.74	±0.65	±0.95
[10.00	15.00]	2.02	±0.17	±0.24

Table C-2a: Differential cross sections in bins of E_T^γ for inclusive prompt photon production in the phase space specified in the table 1.4.

E_T^γ [GeV]		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[6.00	7.00]	$0.93^{+0.02}_{-0.05}$	1.04	0.92 ± 0.00	+4%	−0%
[7.00	8.50]	$0.92^{+0.02}_{-0.04}$	1.04	0.95 ± 0.01	+7%	−0%
[8.50	10.00]	$0.94^{+0.01}_{-0.04}$	1.04	0.97 ± 0.01	+4%	−0%
[10.00	15.00]	$0.95^{+0.02}_{-0.03}$	1.03	0.99 ± 0.01	+9%	−0%

Table C-2b: Corrections applied to the calculations in bins of E_T^γ.

η^γ		E_T^γ [GeV]		$\mathrm{d}\sigma^2/\mathrm{d}E_T^\gamma\mathrm{d}\eta^\gamma$ [pb/[GeV]]	uncorr.	corr.
[−1.00	−0.57]	[6.00	7.00]	9.24	±0.93	±1.06
		[7.00	8.50]	3.75	±0.57	±0.44
		[8.50	10.00]	1.43	±0.41	±0.24
		[10.00	15.00]	0.27	±0.10	±0.07
[−0.57	0.20]	[6.00	7.00]	9.19	±1.04	±0.59
		[7.00	8.50]	5.02	±0.59	±0.48
		[8.50	10.00]	2.29	±0.46	±0.19
		[10.00	15.00]	0.76	±0.12	±0.06
[0.20	0.94]	[6.00	7.00]	10.90	±0.86	±0.78
		[7.00	8.50]	5.15	±0.43	±0.52
		[8.50	10.00]	3.28	±0.34	±0.24
		[10.00	15.00]	0.83	±0.10	±0.07
[0.94	1.42]	[6.00	7.00]	7.68	±1.02	±1.18
		[7.00	8.50]	3.31	±0.51	±0.51
		[8.50	10.00]	2.27	±0.44	±0.38
		[10.00	15.00]	0.66	±0.12	±0.09
[1.42	2.40]	[6.00	7.00]	4.54	±0.90	±1.29
		[7.00	8.50]	2.12	±0.41	±0.72
		[8.50	10.00]	0.86	±0.31	±0.35
		[10.00	15.00]	0.40	±0.08	±0.09

Table C-3a: Double differential cross sections in bins of E_T^γ and η^γ for inclusive prompt photon production in the phase space specified in the table 1.4.

η^γ		E_T^γ [GeV]		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
$[-1.00$	$-0.57]$	[6.00	7.00]	$0.97^{+0.01}_{-0.04}$	1.02	1.00 ± 0.02	$+1\%$	-6%
		[7.00	8.50]	$0.97^{+0.01}_{-0.04}$	1.02	1.01 ± 0.02	$+4\%$	-0%
		[8.50	10.00]	$0.97^{+0.00}_{-0.01}$	1.02	1.01 ± 0.02	$+9\%$	-0%
		[10.00	15.00]	$0.98^{+0.00}_{-0.03}$	1.02	1.01 ± 0.01	$+12\%$	-0%
$[-0.57$	$0.20]$	[6.00	7.00]	$0.94^{+0.02}_{-0.04}$	1.03	0.96 ± 0.00	$+5\%$	-0%
		[7.00	8.50]	$0.94^{+0.02}_{-0.02}$	1.02	0.99 ± 0.01	$+6\%$	-0%
		[8.50	10.00]	$0.94^{+0.02}_{-0.03}$	1.02	1.00 ± 0.01	$+4\%$	-0%
		[10.00	15.00]	$0.98^{+0.01}_{-0.03}$	1.02	1.01 ± 0.01	$+9\%$	-0%
$[0.20$	$0.94]$	[6.00	7.00]	$0.92^{+0.01}_{-0.05}$	1.04	0.92 ± 0.00	$+6\%$	-0%
		[7.00	8.50]	$0.92^{+0.02}_{-0.03}$	1.03	0.95 ± 0.00	$+7\%$	-0%
		[8.50	10.00]	$0.95^{+0.01}_{-0.04}$	1.03	0.98 ± 0.00	$+6\%$	-0%
		[10.00	15.00]	$0.94^{+0.02}_{-0.03}$	1.02	1.00 ± 0.01	$+7\%$	-0%
$[0.94$	$1.42]$	[6.00	7.00]	$0.92^{+0.04}_{-0.04}$	1.06	0.86 ± 0.01	$+5\%$	-0%
		[7.00	8.50]	$0.91^{+0.02}_{-0.05}$	1.05	0.91 ± 0.01	$+10\%$	-0%
		[8.50	10.00]	$0.92^{+0.03}_{-0.04}$	1.05	0.95 ± 0.00	$+6\%$	-0%
		[10.00	15.00]	$0.94^{+0.03}_{-0.03}$	1.04	0.98 ± 0.01	$+10\%$	-0%
$[1.42$	$2.40]$	[6.00	7.00]	$0.91^{+0.02}_{-0.06}$	1.08	0.82 ± 0.00	$+5\%$	-0%
		[7.00	8.50]	$0.89^{+0.02}_{-0.05}$	1.08	0.87 ± 0.01	$+9\%$	-0%
		[8.50	10.00]	$0.91^{+0.02}_{-0.06}$	1.07	0.92 ± 0.02	$+1\%$	-0%
		[10.00	15.00]	$0.92^{+0.03}_{-0.03}$	1.06	0.97 ± 0.02	$+10\%$	-1%

Table C-3b: Corrections applied to the calculations in bins of $E_T^\gamma - \eta^\gamma$.

η^γ		$\mathrm{d}\sigma/\mathrm{d}\eta^\gamma$	uncorr.	corr.
			[pb]	
$[-1.00$	$-0.57]$	14.79	± 1.25	± 1.70
$[-0.57$	$0.20]$	18.57	± 1.47	± 1.75
$[0.20$	$0.94]$	21.12	± 1.21	± 1.77
$[0.94$	$1.42]$	13.88	± 3.15	± 2.12
$[1.42$	$2.40]$	7.31	± 2.73	± 1.54

Table C-4a: Differential cross sections in bins of η^γ for the exclusive prompt photon production in the phase space specified in the table 1.4.

η^γ		f^{MPI}_{theor}	f^{isol}_{theor}	f^{had}_{theor}	δ^{PDF}	
$[-1.00$	$-0.57]$	$0.98^{+0.00}_{-0.04}$	1.02	0.94 ± 0.06	$+4\%$	-0%
$[-0.57$	$0.20]$	$0.95^{+0.02}_{-0.03}$	1.02	0.94 ± 0.04	$+7\%$	-0%
$[0.20$	$0.94]$	$0.93^{+0.02}_{-0.04}$	1.02	0.92 ± 0.03	$+8\%$	-0%
$[0.94$	$1.42]$	$0.93^{+0.03}_{-0.04}$	1.04	0.90 ± 0.02	$+10\%$	-0%
$[1.42$	$2.40]$	$0.91^{+0.02}_{-0.05}$	1.06	0.88 ± 0.03	$+8\%$	-0%

Table C-4b: Corrections applied to the calculations in bins of η^γ.

E_T^γ [GeV]		$\mathrm{d}\sigma/\mathrm{d}E_T^\gamma$ [pb/[GeV]]	uncorr.	corr.
[6.00	7.00]	18.53	$\pm$1.61	$\pm$2.21
[7.00	8.50]	9.93	$\pm$1.06	$\pm$1.39
[8.50	10.00]	5.50	$\pm$0.45	$\pm$0.73
[10.00	15.00]	1.68	$\pm$0.23	$\pm$0.16

Table C-5a: Differential cross sections in bins of E_T^γ for the exclusive prompt photon production in the phase space specified in the table 1.4.

E_T^γ [GeV]		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[6.00	7.00]	$0.94^{+0.02}_{-0.05}$	1.03	0.85 ± 0.05	+8%	−0%
[7.00	8.50]	$0.92^{+0.02}_{-0.04}$	1.03	0.92 ± 0.04	+8%	−0%
[8.50	10.00]	$0.94^{+0.02}_{-0.04}$	1.03	0.96 ± 0.02	+5%	−0%
[10.00	15.00]	$0.95^{+0.02}_{-0.03}$	1.03	0.99 ± 0.01	+8%	−0%

Table C-5b: Corrections applied to the calculations in bins of E_T^γ.

η^{jet}		$\mathrm{d}\sigma/\mathrm{d}\eta^{jet}$	uncorr. [pb]	corr.
$[-1.30$	$-0.40]$	7.10	± 0.69	± 0.97
$[-0.40$	$0.50]$	14.87	± 0.85	± 2.10
$[0.50$	$1.40]$	18.73	± 1.07	± 2.74
$[1.40$	$2.30]$	15.35	± 1.16	± 2.15

Table C-6a: Differential cross sections in bins of η^{jet} for the exclusive prompt photon production in the phase space specified in the table 1.4.

η^{jet}		f^{MPI}_{theor}	f^{isol}_{theor}	f^{had}_{theor}	δ^{PDF}	
$[-1.30$	$-0.40]$	$0.97^{+0.01}_{-0.02}$	1.02	0.81 ± 0.03	$+12\%$	-0%
$[-0.40$	$0.50]$	$0.95^{+0.01}_{-0.03}$	1.02	0.89 ± 0.03	$+8\%$	-0%
$[0.50$	$1.40]$	$0.92^{+0.02}_{-0.04}$	1.03	0.97 ± 0.04	$+5\%$	-0%
$[1.40$	$2.30]$	$0.91^{+0.03}_{-0.06}$	1.05	0.99 ± 0.02	$+6\%$	-0%

Table C-6b: Corrections applied to the calculations in bins of η^{jet}.

E_T^{jet} [GeV]		$\mathrm{d}\sigma/\mathrm{d}E_T^{jet}$ [pb/[GeV]]	uncorr.	corr.
[4.50	6.20]	6.60	±0.74	±1.42
[6.20	8.00]	6.93	±1.08	±0.84
[8.00	10.00]	6.15	±0.78	±0.65
[10.00	15.00]	1.88	±0.33	±0.17

Table C-7a: Differential cross sections in bins of E_T^{jet} for the exclusive prompt photon production in the phase space specified in the table 1.4.

E_T^{jet} [GeV]		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[4.50	6.20]	$0.88^{+0.03}_{-0.05}$	1.02	0.95 ± 0.03	+4%	−0%
[6.20	8.00]	$0.92^{+0.02}_{-0.04}$	1.02	0.88 ± 0.04	+9%	−0%
[8.00	10.00]	$0.97^{+0.01}_{-0.04}$	1.03	0.91 ± 0.03	+9%	−0%
[10.00	15.00]	$1.00^{+0.01}_{-0.03}$	1.04	0.94 ± 0.03	+8%	−0%

Table C-7b: Corrections applied to the calculations in bins of E_T^{jet}.

x_γ^{LO}		$\mathrm{d}\sigma/\mathrm{d}x_\gamma^{LO}$	uncorr. [pb]	corr.
[0.00	0.50]	23.03	±3.31	±3.80
[0.50	0.70]	44.10	±9.64	±8.11
[0.70	0.90]	70.33	±12.25	±10.90
[0.90	1.10]	75.89	±9.83	±8.41
[1.10	1.40]	3.33	±2.76	±1.07

Table C-8a: Differential cross sections in bins of x_γ^{LO} for the exclusive prompt photon production in the phase space specified in the table 1.4.

x_γ^{LO}		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[0.00	0.50]	$0.83^{+0.04}_{-0.08}$	1.09	0.86 ± 0.02	+2%	−0%
[0.50	0.70]	$0.87^{+0.04}_{-0.07}$	1.06	0.97 ± 0.05	+0%	−4%
[0.70	0.90]	$0.96^{+0.02}_{-0.04}$	1.03	1.27 ± 0.04	+0%	−5%
[0.90	1.10]	$1.00^{+0.00}_{-0.01}$	1.00	0.82 ± 0.04	+14%	−0%
[1.10	1.40]	$0.99^{+0.00}_{-0.01}$	1.00	0.44 ± 0.09	+10%	−0%

Table C-8b: Corrections applied to the calculations in bins of x_γ^{LO}.

x_p^{LO}		$\mathrm{d}\sigma/\mathrm{d}x_p^{LO}$	uncorr. [pb]	corr.
[0.00	0.01]	1257.22	±91.07	±101.61
[0.01	0.03]	1324.65	±64.87	±149.67
[0.03	0.04]	697.78	±90.08	±133.83
[0.04	0.06]	341.36	±46.72	±66.28

Table C-9a: Differential cross sections in bins of x_p^{LO} for the exclusive prompt photon production in the phase space specified in the table 1.4.

x_p^{LO}		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[0.00	0.01]	$0.96^{+0.01}_{-0.03}$	1.02	0.86 ± 0.04	+10%	−0%
[0.01	0.03]	$0.94^{+0.02}_{-0.04}$	1.02	0.92 ± 0.04	+6%	−0%
[0.03	0.04]	$0.92^{+0.02}_{-0.04}$	1.04	0.95 ± 0.03	+6%	−0%
[0.04	0.06]	$0.91^{+0.03}_{-0.05}$	1.06	0.96 ± 0.02	+9%	−0%

Table C-9b: Corrections applied to the calculations in bins of x_p^{LO}.

x_γ^{LO}		$p_\perp$ [GeV]		$\mathrm{d}\sigma^2/\mathrm{d}p_\perp\mathrm{d}x_\gamma^{LO}$ [pb/[GeV]]	uncorr.	corr.
[0.8	1.2]	[0.00	2.00]	7.75	±0.50	±0.32
		[2.00	4.00]	3.96	±0.43	±0.53
		[4.00	6.00]	2.16	±0.55	±0.56
		[6.00	8.00]	0.60	±0.53	±0.36
[0.	0.8]	[0.00	2.00]	7.14	±0.48	±1.04
		[2.00	4.00]	4.52	±0.49	±0.80
		[4.00	6.00]	2.91	±0.40	±0.56
		[6.00	8.00]	2.35	±0.44	±0.37

Table C-10a: Differential cross sections in bins of $p_\perp$ for the exclusive prompt photon production in the phase space specified in the table 1.5.

x_γ^{LO}		$p_\perp$ [GeV]		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[0.8	1.2]	[0.00	2.00]	$0.99^{+0.00}_{-0.01}$	1.00	0.90 ± 0.04	+14%	−0%
		[2.00	4.00]	$1.00^{+0.00}_{-0.01}$	1.01	0.86 ± 0.00	+0%	−5%
		[4.00	6.00]	$1.01^{+0.00}_{-0.03}$	1.01	0.83 ± 0.07	+6%	−0%
		[6.00	8.00]	$0.97^{+0.02}_{-0.00}$	1.01	0.83 ± 0.08	+14%	−2%
[0.	0.8]	[0.00	2.00]	$0.79^{+0.04}_{-0.08}$	1.07	0.99 ± 0.07	+0%	−3%
		[2.00	4.00]	$0.88^{+0.03}_{-0.07}$	1.07	0.98 ± 0.01	+1%	−1%
		[4.00	6.00]	$0.97^{+0.03}_{-0.04}$	1.06	0.93 ± 0.01	+0%	−4%
		[6.00	8.00]	$1.07^{+0.01}_{-0.06}$	1.06	0.94 ± 0.08	+2%	−2%

Table C-10b: Corrections applied to the calculations in bins of $p_\perp$.

x_γ^{LO}		$\Delta\Phi$ [deg]		$d\sigma^2/d\Delta\Phi dx_\gamma^{LO}$ [pb/[deg]]	uncorr.	corr.
[0.8	1.2]	[130.00	150.00]	0.19	±0.04	±0.01
		[150.00	165.00]	0.46	±0.06	±0.06
		[165.00	172.00]	0.89	±0.14	±0.07
		[172.00	180.00]	1.38	±0.04	±0.17
[0.	0.8]	[130.00	150.00]	0.27	±0.03	±0.04
		[150.00	165.00]	0.52	±0.06	±0.08
		[165.00	172.00]	0.91	±0.14	±0.18
		[172.00	180.00]	1.21	±0.11	±0.15

Table C-11a: Differential cross sections in bins of $\Delta\Phi$ for the exclusive prompt photon production in the phase space specified in the table 1.5.

x_γ^{LO}		$\Delta\Phi$ [deg]		f_{theor}^{MPI}	f_{theor}^{isol}	f_{theor}^{had}	δ^{PDF}	
[0.8	1.2]	[130.00	150.00]	$1.04^{+0.00}_{-0.04}$	1.01	0.78 ± 0.03	+0%	−5%
		[150.00	165.00]	$0.99^{+0.01}_{-0.01}$	1.01	0.84 ± 0.01	+0%	−3%
		[165.00	172.00]	$0.99^{+0.00}_{-0.01}$	1.01	0.93 ± 0.00	+7%	−0%
		[172.00	180.00]	$0.99^{+0.01}_{-0.01}$	1.00	0.89 ± 0.05	+16%	−0%
[0.	0.8]	[130.00	150.00]	$0.95^{+0.03}_{-0.04}$	1.06	0.94 ± 0.01	+0%	−3%
		[150.00	165.00]	$0.87^{+0.03}_{-0.08}$	1.07	0.98 ± 0.01	+0%	−2%
		[165.00	172.00]	$0.80^{+0.04}_{-0.07}$	1.07	0.98 ± 0.06	+0%	−2%
		[172.00	180.00]	$0.79^{+0.05}_{-0.09}$	1.07	1.00 ± 0.08	+1%	−3%

Table C-11b: Corrections applied to the calculations in bins of $\Delta\Phi$.

List of Figures

List of Tables

References

[1] C. Amsler *et al.*, "Review of particle physics," *Phys. Lett.*, vol. B667, p. 1, 2008.

[2] D. Sunar, *Measurement of $K^{*\pm}$ production in deep inelastic ep scattering with the H1 detector at HERA*. PhD thesis, Universiteit Antwerpen, 2009.

[3] A. I. Abazov *et al.*, "Search for neutrinos from sun using the reaction Ga-71 (electron-neutrino e-) Ge-71," *Phys. Rev. Lett.*, vol. 67, pp. 3332–3335, 1991.

[4] P. Anselmann *et al.*, "Solar neutrinos observed by GALLEX at Gran Sasso.," *Phys. Lett.*, vol. B285, pp. 376–389, 1992.

[5] Y. Fukuda *et al.*, "Solar neutrino data covering solar cycle 22," *Phys. Rev. Lett.*, vol. 77, pp. 1683–1686, 1996.

[6] T. Araki *et al.*, "Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion," *Phys. Rev. Lett.*, vol. 94, p. 081801, 2005, hep-ex/0406035.

[7] J. Hosaka *et al.*, "Solar neutrino measurements in Super-Kamiokande-I," *Phys. Rev.*, vol. D73, p. 112001, 2006, hep-ex/0508053.

[8] D. J. Griffiths, *Introduction to Elementary Particles*. New York, USA, Wiley, 1987.

[9] F. Englert and R. Brout, "Broken symmetry and themass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–322, 1964.

[10] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.

[11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

[12] C. Adloff *et al.*, "Measurement and QCD analysis of neutral and charged current cross sections at HERA," *Eur. Phys. J.*, vol. C30, pp. 1–32, 2003, hep-ex/0304003.

[13] F. D. Aaron *et al.*, "Jet production in ep collisions at high Q^2 and determination of α_s," 2009, 0904.3870.

[14] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972.

[15] V. N. Gribov and L. N. Lipatov, "e+ e- pair annihilation and deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.*, vol. 15, pp. 675–684, 1972.

[16] L. N. Lipatov, "The parton model and perturbation theory," *Sov. J. Nucl. Phys.*, vol. 20, pp. 94–102, 1975.

[17] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language," *Nucl. Phys.*, vol. B126, p. 298, 1977.

[18] Y. L. Dokshitzer, "Calculation of the structure functions for deep inelastic scattering and e+ e- annihilation by perturbation theory in quantum chromodynamics," *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977.

[19] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, "The pomeranchuk singularity in nonabelian gauge theories," *Sov. Phys. JETP*, vol. 45, pp. 199–204, 1977.

[20] I. I. Balitsky and L. N. Lipatov, "The pomeranchuk singularity in quantum chromodynamics," *Sov. J. Nucl. Phys.*, vol. 28, pp. 822–829, 1978.

[21] S. Catani, M. Ciafaloni, and F. Hautmann, "High-energy factorization and small x heavy flavor production," *Nucl. Phys.*, vol. B366, pp. 135–188, 1991.

[22] M. Ciafaloni, "Coherence Effects in Initial Jets at Small q**2 / s," *Nucl. Phys.*, vol. B296, p. 49, 1988.

[23] S. Catani, F. Fiorani, and G. Marchesini, "QCD Coherence in Initial State Radiation," *Phys. Lett.*, vol. B234, p. 339, 1990.

[24] F. D. Aaron *et al.*, "A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA," 2009, 0904.3513.

[25] C. Adloff *et al.*, "Measurement and QCD analysis of neutral and charged current cross sections at HERA," *Eur. Phys. J.*, vol. C30, pp. 1–32, 2003, hep-ex/0304003.

[26] H. L. Lai *et al.*, "Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions," *Eur. Phys. J.*, vol. C12, pp. 375–392, 2000, hep-ph/9903282.

[27] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Physical Gluons and High E(T) jets," *Phys. Lett.*, vol. B604, pp. 61–68, 2004, hep-ph/0410230.

[28] C. F. von Weizsacker, "Radiation emitted in collisions of very fast electrons," *Z. Phys.*, vol. 88, pp. 612–625, 1934.

[29] E. J. Williams, "Nature of the high-energy particles of penetrating radiation and status of ionization and radiation formulae," *Phys. Rev.*, vol. 45, pp. 729–730, 1934.

[30] L. Finke, *Measurement of charm and beauty dijet cross sections in photoproduction.* PhD thesis, Universität Hamburg, 2006.

[31] E. Witten, "Anomalous Cross-Section for Photon - Photon Scattering in Gauge Theories," *Nucl. Phys.*, vol. B120, pp. 189–202, 1977.

[32] R. Nisius, "The photon structure from deep inelastic electron photon scattering," *Phys. Rept.*, vol. 332, pp. 165–317, 2000, hep-ex/9912049.

[33] W. T. Giele and E. W. N. Glover, "Higher order corrections to jet cross-sections in e+ e- annihilation," *Phys. Rev.*, vol. D46, pp. 1980–2010, 1992.

[34] A. Gehrmann-De Ridder, T. Gehrmann, and E. Poulsen, "Measuring the photon fragmentation function at HERA," *Eur. Phys. J.*, vol. C47, pp. 395–411, 2006, hep-ph/0604030.

[35] C. A.Aktas *et al.*, "Measurement of Prompt Photon Cross Sections in Photoproduction at HERA," *Eur. Phys. J.*, vol. C38, p. 437, 2005.

[36] J. Breitweg *et al.*, "Measurement of inclusive prompt photon photoproduction at HERA," *Phys. Lett.*, vol. B472, pp. 175–188, 2000, hep-ex/9910045.

[37] S. Chekanov *et al.*, "Study of the effective transverse momentum of partons in the proton using prompt photons in photoproduction at HERA," *Phys. Lett.*, vol. B511, pp. 19–32, 2001, hep-ex/0104001.

[38] S. Chekanov *et al.*, "Measurement of prompt photons with associated jets in photoproduction at HERA," *Eur. Phys. J.*, vol. C49, pp. 511–522, 2007, hep-ex/0608028.

[39] M. Fontannaz, J. P. Guillet, and G. Heinrich, "Isolated Prompt Photon Photoproduction at NLO," *Eur. Phys. J.*, vol. C21, pp. 303–312, 2001, hep-ph/0105121.

[40] M. Krawczyk and A. Zembrzuski, "Photoproduction of the isolated photon at DESY HERA in next-to-leading order QCD," *Phys. Rev.*, vol. D64, p. 114017, 2001, hep-ph/0105166.

[41] S. Chekanov *et al.*, "Observation of isolated high-E(T) photons in deep inelastic scattering," *Phys. Lett.*, vol. B595, pp. 86–100, 2004, hep-ex/0402019.

[42] M. Forrest, "Prompt Photon Production in DIS," Prepared for 17th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2009), Madrid, Spain, 26-30 Apr 2009.

[43] F. D. Aaron *et al.*, "Measurement of Isolated Photon Production in Deep- Inelastic Scattering at HERA," *Eur. Phys. J.*, vol. C54, pp. 371–387, 2008, 0711.4578.

[44] P. Aurenche, M. Fontannaz, J.-P. Guillet, E. Pilon, and M. Werlen, "A New critical study of photon production in hadronic collisions," *Phys. Rev.*, vol. D73, p. 094007, 2006, hep-ph/0602133.

[45] H.-L. Lai and H.-n. Li, "Origin of the k_T smearing in direct photon production," *Phys. Rev.*, vol. D58, p. 114020, 1998, hep-ph/9802414.

[46] P. Aurenche *et al.*, "A critical phenomenological study of inclusive photon production in hadronic collisions," *Eur. Phys. J.*, vol. C9, pp. 107–119, 1999, hep-ph/9811382.

[47] C. D. Silberberg, *Measurement of the inclusive ssolated prompt photon production cross section at the Tevatron using the CDF detector*. PhD thesis, Universitat Autònoma de Barcelona, 2009.

[48] L. E. Gordon and W. Vogelsang, "Isolated prompt photon production at HERA," *Phys. Rev.*, vol. D52, pp. 58–67, 1995.

[49] M. Fontannaz, J. P. Guillet, and G. Heinrich, "Is a large intrinsic k(T) needed to describe photon + jet photoproduction at HERA?," *Eur. Phys. J.*, vol. C22, pp. 303–315, 2001, hep-ph/0107262.

[50] P. Aurenche, J. P. Guillet, and M. Fontannaz, "Parton Distributions in the Photon," *Z. Phys.*, vol. C64, pp. 621–630, 1994, hep-ph/9406382.

[51] M. Fontannaz, J. P. Guillet, and G. Heinrich, "A NLO calculation of the hadron-jet cross section in photoproduction reactions," *Eur. Phys. J.*, vol. C26, pp. 209–218, 2002, hep-ph/0206202.

[52] P. Aurenche, "Inclusive hard processes in photon induced reactions," 1997, hep-ph/9706386.

[53] S. Frixione and G. Ridolfi, "Jet photoproduction at HERA," *Nucl. Phys.*, vol. B507, pp. 315–333, 1997, hep-ph/9707345.

[54] S. Caron, *Jets in Photoproduction at HERA*. PhD thesis, RWTH Aachen, 2002.

[55] M. Bengtsson and T. Sjostrand, "Parton Showers in Leptoproduction Events," *Z. Phys.*, vol. C37, p. 465, 1988.

[56] G. Gustafson, "Dual Description of a Confined Color Field," *Phys. Lett.*, vol. B175, p. 453, 1986.

[57] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, "Parton Fragmentation and String Dynamics," *Phys. Rept.*, vol. 97, pp. 31–145, 1983.

[58] B. Andersson, G. Gustafson, and B. Soderberg, "A General Model for Jet Fragmentation," *Z. Phys.*, vol. C20, p. 317, 1983.

[59] G. C. Fox and S. Wolfram, "A Model for Parton Showers in QCD," *Nucl. Phys.*, vol. B168, p. 285, 1980.

[60] V. Shekelyan, "Simulation and Reconstruction in H1 Liquid Argon Calorimetry," *H1-Internal Note, H1-0493-288, DESY*, 1993.

[61] *"GEANT - Detector Description and Simulation Tool"*. CERN Program Library Long Writeup W5013.

[62] H. C. Fesefeldt, "Simulation of hadronic showers, physics and applications," Tech. Rep. PITHA 85-02, III Physikalisches Institut, RWTHAachen Physikzentrum, 5100 Aachen, Germany, 1985.

[63] P. A. Aarnio *et al.*, "Fluka users guide," Tech. Rep. TIS-RP-190, CERN, 1987, 1990.

[64] S. Peters, *Die Parametriserte Simulation Elektromagnetischer Schauer*. PhD thesis, MPI Münich, 1992.

[65] M. Ellerbrock *et al.*, "Improved Description of Electromagnetic Showers in H1FAST," *H1-Internal Note, H1-01/05-617, DESY*, 2005.

[66] C. Schmitz, "Isolated Photon Production in Deep-Inelastic Scattering at HERA - Ph.D. thesis," 2007.

[67] I. Abt *et al.*, "The Tracking, Calorimeter and Muon Detectors of the H1 Experiment at HERA ," *Nucl. Instrum. Meth.*, vol. A386, pp. 348–396, 1997.

[68] T. Sjostrand, L. Lonnblad, and S. Mrenna, "PYTHIA 6.2: Physics and Manual," 2001, hep-ph/0108264.

[69] T. Sjostrand, "High-Energy Physics Event Generation with PYTHIA 5.7 and JETSET 7.4," *Comput. Phys. Commun.*, vol. 82, pp. 74–90, 1994.

[70] G. Corcella *et al.*, "HERWIG 6.5 Release Note," 2002, hep-ph/0210213.

[71] H. Jung, "Hard Diffractive Scattering in High-Energy *ep* collisions and the Monte Carlo generator RAPGAP," *Comp. Phys. Commun.*, vol. 86, pp. 147–161, 1995.

[72] H. Spiesberger, *"HERACLES and DJANGOH: Event Generation of ep Interactions at HERA Including Radiative Processes"*. User manual.

[73] J. Pumplin *et al.*, "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis," *JHEP*, vol. 07, p. 012, 2002, hep-ph/0201195.

[74] G. A. Schuler and T. Sjostrand, "Parton Distributions of the Virtual Photon," *Phys. Lett.*, vol. B376, pp. 193–200, 1996, hep-ph/9601282.

[75] M. Fontannaz and G. Heinrich, "Isolated Photon + Jet Photoproduction as a Tool to Constrain the Gluon Distribution in the Proton and the Photon," *Eur. Phys. J.*, vol. C34, pp. 191–199, 2004, hep-ph/0312009.

[76] L. V. Gribov, E. M. Levin, and M. G. Ryskin, "Semihard Processes in QCD," *Phys. Rept.*, vol. 100, pp. 1–150, 1983.

[77] S. Catani, M. Ciafaloni, and F. Hautmann, "High-Energy Factorization and Small x Heavy Flavor Production," *Nucl. Phys.*, vol. B366, pp. 135–188, 1991.

[78] J. R. Andersen *et al.*, "Small x Phenomenology: Summary and Status 2002," *Eur. Phys. J.*, vol. C35, pp. 67–98, 2004, hep-ph/0312333.

[79] J. R. Andersen *et al.*, "Small x Phenomenology: Summary of the 3rd Lund Small x Workshop in 2004," *Eur. Phys. J.*, vol. C48, pp. 53–105, 2006, hep-ph/0604189.

[80] A. V. Lipatov and N. P. Zotov, "Prompt Photon Photoproduction at HERA in the k(T)- Factorization Approach," *Phys. Rev.*, vol. D72, p. 054002, 2005, hep-ph/0506044.

[81] M. A. Kimber, A. D. Martin, and M. G. Ryskin, "Unintegrated Parton Distributions," *Phys. Rev.*, vol. D63, p. 114027, 2001, hep-ph/0101348.

[82] M. Gluck, E. Reya, and A. Vogt, "Photonic Parton Distributions," *Phys. Rev.*, vol. D46, pp. 1973–1979, 1992.

[83] M. Gluck, E. Reya, and A. Vogt, "Dynamical Parton Distributions of the Proton and Small x Physics," *Z. Phys.*, vol. C67, pp. 433–448, 1995.

[84] P. Schmueser, "The electron proton colliding beam facility HERA," *Nucl. Instrum. Meth.*, vol. A235, pp. 201–208, 1985.

[85] I. Abt *et al.*, "The H1 Detector at HERA," *Nucl. Instrum. Meth.*, vol. A386, pp. 310–347, 1997.

[86] "The ZEUS detector: Status report 1993," ZEUS-STATUS-REPT-1993.

[87] K. Ackerstaff *et al.*, "HERMES spectrometer," *Nucl. Instrum. Meth.*, vol. A417, pp. 230–265, 1998, hep-ex/9806008.

[88] P. Krizan, R. Mankel, D. Ressing, S. Shuvalov, and M. Spahn, "HERA-B, an experiment to study CP violation at the HERA proton ring using an internal target," *Nucl. Instrum. Meth.*, vol. A351, pp. 111–131, 1994.

[89] J. M. J. K. Hirata, "The luminosity upgrade of hera," *ICFA Beam Dynamics Newsletter, No. 24*, 2001.

[90] J. Burger *et al.*, "The Central jet chamber of the H1 experiment," *Nucl. Instrum. Meth.*, vol. A279, pp. 217–222, 1989.

[91] J. Becker *et al.*, "A vertex trigger based on cylindrical multiwire proportional chambers," *Nucl. Instrum. Meth.*, vol. A586, pp. 190–203, 2008, physics/0701002.

[92] S. Egli *et al.*, "Calibration of the outer z drift chamber of the HERA H1 experiment," *DESY-ZEUTHEN-94-03*, vol. A283, p. 1949.

[93] S. Burke *et al.*, "Track finding and fitting in the H1 forward track detector," *Nucl. Instrum. Meth.*, vol. A373, pp. 227–260, 1996.

[94] A. Makankin *et al.*, "Backward proportinal chamber at the H1 experiment," *http://www-h1.desy.de/h1det/tracker/bpc/BPCpage.htm.*

[95] W. Eick *et al.*, "Development of the H1 backward silicon strip detector," *Nucl. Instrum. Meth.*, vol. A386, pp. 81–86, 1997.

[96] D. Pitzl *et al.*, "The H1 silicon vertex detector," *Nucl. Instrum. Meth.*, vol. A454, pp. 334–349, 2000, hep-ex/0002044.

[97] W. Braunschweig *et al.*, "A forward silicon tracker for H1," *H1 note, DESY-PRC-99/01*, 1999.

[98] B. Andrieu *et al.*, "The H1 liquid argon calorimeter system," *Nucl. Instrum. Meth.*, vol. A336, pp. 460–498, 1993.

[99] W. M. Yao *et al.*, "Review of particle physics," *J. Phys.*, vol. G33, pp. 1–1232, 2006.

[100] G. Bathow, E. Freytag, M. Koebberling, K. Tesch, and R. Kajikawa, "Measurements of the longitudinal and lateral development of electromagnetic cascades in lead, copper and aluminum at 6 gev," *Nucl. Phys.*, vol. B20, pp. 592–602, 1970.

[101] W. R. Nelson, T. M. Jenkins, R. C. McCall, and J. K. Cobb, "Electron induced cascade showers in copper and lead 1 GeV," *Phys. Rev.*, vol. 149, pp. 201–208, 1966.

[102] B. Andrieu *et al.*, "Beam Tests and Calibration of the H1 Liquid Argon Calorimeter with Electrons," *Nucl. Instrum. Meth.*, vol. A350, pp. 57–72, 1994.

[103] B. Andrieu *et al.*, "Results from Pion Calibration Runs for the H1 Liquid Argon Calorimeter and Comparisons with Simulations," *Nucl. Instrum. Meth.*, vol. A336, pp. 499–509, 1993.

[104] R. D. Appuhn *et al.*, "The H1 lead/scintillating-fibre calorimeter," *Nucl. Instrum. Meth.*, vol. A386, pp. 397–408, 1997.

[105] V. Andreev *et al.*, "Proposal for an upgrade of the H1 luminosity system and its associated electronics for HERA 2000," *H1 note, DESY-PRC-98/05*, 1998.

[106] V. Karimäki, "Fast code to fit circular arcs," *Comput. Phys. Commun.*, vol. 69, pp. 133–141, 1992.

[107] J. Krosberg, *A Measurement of Beauty Production in High Energy Positron-Proton Scattering*. PhD thesis, Universität Zürich, 2002.

[108] P. Bruel, *Recherche d'interactions au-dela du Modele Standard a HERA*. PhD thesis, University Paris, 1998.

[109] S. D. Ellis and D. E. Soper, "Successive Combination Jet Algorithm for Hadron Collisions," *Phys. Rev.*, vol. D48, pp. 3160–3166, 1993, hep-ph/9305266.

[110] M. H. Seymour, "Jets in QCD," *AIP Conf. Proc.*, vol. 357, pp. 568–587, 1996, hep-ph/9506421.

[111] M. zur Nedden, B. Reisert, T. Schörner, "H1 Liquid Argon Trigger: Overview, Simulation and Performance," *H1-Internal Note, H1-04/01-592, DESY*, 2001.

[112] C. Veelken, "H1NonepBgFinder - Rejection of Cosmic Muon and Beam-Halo Events in the H1OO framework," *H1-Internal Note*, vol. H1-09/02-603, DESY, 2002.

[113] E. Chabert, C. Diaconu, S. Kermiche, I. Negri, C. Vallee, "QBGFMAR - an Updated PHAN Package for Cosmic and Halo Muon Topological Rejection in High P_T Physics Analysis," *H1-Internal Note*, vol. H1-11/98-556, DESY, 1998.

[114] F. Jacquet and A. Blondel, "Report from the Study Group on Detectors for Charged Current Events," *Hamburg 1979, Proceedings, Study of an ep Facility for Europe*, Hamburg 1979, 377-414.

[115] J.F.Kenney and E.S.Keeping, "Mathematics of Statistics, NJ: Van Nostrand, 1962,"

[116] A. Hocker *et al.*, "TMVA: Toolkit for multivariate data analysis," *PoS*, vol. ACAT, p. 040, 2007, physics/0703039.

[117] P.Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss," *Machine Learning*, vol. 29, pp. 103–130, 1997.

[118] B. Heinemann, *Measurement of Charged Current and Neutral Current Cross Sections in Positron-Proton Collisions at sqrt(s)=300 GeV.* PhD thesis, University of Hamburg, 1999.

[119] H. Collaboration, "Deeply Virtual Compton Scattering and its Beam Charge Asymmetry in $e^{\pm}$ p Collisions at HERA," 2009, 0907.5289.

[120] M. Peez, B. Portheault, E. Sauvan, "An Energy Flow Algorithm for Hadronic Reconstruction in OO: Hadroo2," *H1-Internal Note*, vol. H1-01/05-616, DESY, 2005.

[121] V. Blobel, "An unfolding method for high energy physics experiments," 2002, hep-ex/0208022.

[122] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework, TUnfold package,"

[123] P. C. Hansen, "The l-curve and its use in the numerical treatment of inverse problems, http://www.imm.dtu.dk/ pch/tr/lcurve.ps," *Advances in Computational Bioengineering.*

[124] G. D'Agostini, "On the use of the covariance matrix to fit correlated data," *Nucl. Instrum. Meth.*, vol. A346, pp. 306–311, 1994.

[125] F. Abe *et al.*, "Measurement of Double Parton Scattering in $\bar{p}p$ Collisions at $\sqrt{s} =$ 1.8 TeV," *Phys. Rev. Lett.*, vol. 79, pp. 584–589, 1997.

[126] L. Marti, *Multiple Parton Interactions in Photoproduction at HERA/H1.* PhD thesis, University of Hamburg, 2009.

[127] C. Gwenlan, "Multijets in Photoproduction at HERA," *Acta Phys. Polon.*, vol. B33, pp. 3123–3128, 2002.

[128] S. Aid *et al.*, "Jets and Energy Flow in Photon - Proton Collisions at HERA," *Z. Phys.*, vol. C70, pp. 17–30, 1996, hep-ex/9511012.

[129] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, "Multiparton Interactions in Photoproduction at HERA," *Z. Phys.*, vol. C72, pp. 637–646, 1996, hep-ph/9601371.

[130] T. Sjostrand and M. van Zijl, "A Multiple Interaction Model for the Event Structure in Hadron Collisions," *Phys. Rev.*, vol. D36, p. 2019, 1987.

[131] S. Magill, ed., *Optimization of Jet Algorithm Inputs in the ZEUS Detector, Proceedings of the IX International Conference on Calorimetry in High Energy Physics*, vol. XXI, (Annecy Le Vieux Cedex, France), October 2000.

[132] B. T.Carli *Nucl.Instrum.Meth.A501*, p. 576, 2003. hep-ex/0211019.

[133] R. Sedgewick, *Algorithms in C++*, ch. 26. Addison Wesley, 1992.

[134] B. V. Dasarathy, *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. Los Alamitos, Calif. : IEEE Computer Society Press, 1991. ISBN 0-8186-8930-7.

[135] R.A.Fisher, "The use of multiple measurements in taxoconic problems," *Annals Eugenics*, vol. 7, p. 179, 1936.

[136] C.Bishop, *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press, 1995.

[137] E. Bishop, "A generalization of the stone-weierstrass theorem," *Pacific J. Math.*, vol. 11, no. 3, pp. 771–783, 1961.

[138] P.Werbos, *The Roots of Backpropagation*. New York: Wiley, 1994.

[139] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, p. 1, 1998.

[140] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting.," *Journal of Computer and System Science*, vol. 55, pp. 119–139, 1997.